AFOSR 70 -28 09 TR

AD716591
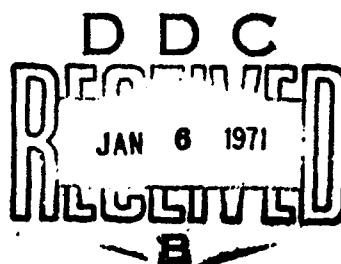
STATISTICS LABORATORY

MATHEMATICS DEPARTMENT

THE OHIO STATE UNIVERSITY

COLUMBUS, OHIO

D D C

RECEIVED

JAN 6 1971

B

47

PRINCIPAL COMPONENT AN/LYSIS
OF TIME SERIES [1,2]

J. Richard Stewart

Division of Statistics

The Ohio State University

September, 1970

ACKNOWLEDGMENTS

INTRODUCTION

A situation which sometimes arises is that
in which a time series y(t), t = 0,1,... is to be
modelled as a function of another time series x(t),
t = 0,1,... (or more generally as a function of k series
$x_1(t),...,x_k(t)$; t = 0,1,...). It is often the case
that for any time t, y(t) will depend not only on x(t),
but also on x(t - 1), x(t - 2),...,x(t - r) for some in-
teger r ≥ 0, and sometimes on the complete past history
of x(t). When all of the series are second order sta-
tionary and a large number of observations is available,
the powerful methods of spectral analysis may be employ-
ed to estimate various types of linear models. On the
other hand, when relatively few observations are avail-
able and the series involved may be nonstationary, at
least to the extent of having trending means, these
methods may not be appropriate, and yet it may be desir-
able to have some sort of frequency decomposition of the
series. (For instance, it may be of interest to model
trends and stationary parts of the series separately.)

1

The purpose of this study is to investigate some of the properties of a method whereby $y(t)$ is modelled not directly as a linear function of $x(t)$, $x(t-1),\ldots,x(t-r)$, but rather as a linear function of some of the principal components of $x(t),x(t-1)$, $\ldots,x(t-r)$. Chapter I, formulated in terms of a set of random variables $y,x_1,\ldots,x_p$, investigates the co-variance structure among these variables which will yield the result that the first $m < p$ principal components of $x_1,\ldots,x_p$ provide a better predictor of y than does any subset $x_{i_1},\ldots,x_{i_m}$.

In chapter II, principal component processes of $x(t),x(t-1),\ldots,x(t-n)$ are defined and their properties noted. It is shown that when $x(t)$ is one of the usual types of second order stationary processes, the principal component processes effect a partial frequency decomposition of $x(t)$.

In chapter III, it is shown that for certain processes containing deterministic components, certain of the principal component processes tend to filter out these deterministic components.

An example of the behavior of the principal
component processes of an artificially generated process
is shown in the appendix together with the principal
components of an autoregressive process, some of the
transfer functions corresponding to the eigenvectors
of a moving average process, and the first two principal
component processes derived from a series representing
quarterly total U.S. personal income.

CHAPTER I

REGRESSION ON PRINCIPAL COMPONENTS

Let X be the p component random vector $\begin{bmatrix} x_1 \\ \cdot \\ \cdot \\ \cdot \\ x_p \end{bmatrix}$

with covariance matrix Σ. It will be assumed that $EX = 0$, since the mean vector does not enter into the following discussion. There exists an orthogonal linear transformation

$$Z = \phi'X$$

such that the covariance matrix of Z,

$$E(ZZ') = \Lambda = \begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ & & \ddots \\ 0 & & \lambda_p \end{bmatrix},$$

where $\lambda_1 \geq \ldots \geq \lambda_p$ are the roots (eigenvalues) of the characteristic equation $|\Sigma - \lambda I| = 0$, and $\phi = (\phi_1 \ldots \phi_p)$, with ϕ_j the eigenvector corresponding to λ_j. The rth component of Z, $Z_r = \phi_r'X$ is called the rth principal component of X and has maximum variance of all normalized ($\phi_r' \phi_r = 1$) linear combinations uncorrelated with $Z_1, \ldots, Z_{r-1}$. In particular Z_1 the first principal component has maximum variance of any normalized linear

4

combination. Moreover,

$$\sum_{i=1}^{p} E(Z_i)^2 = \sum_{i=1}^{p} \lambda_i = tr\Lambda = tr\Sigma = \sum_{i=1}^{p} E(x_i)^2 \quad (1.0)$$

Principal components have found application especially in exploratory studies where the number p of variables under consideration may be large. It may happen that a few of the principal components explain most of the variation (as shown by (1.0)) of the original p variables. Another criterion often applied to the use of principal components is that they should have some sort of "reasonable" interpretation in their own right.

When the variables $x_1,...,x_p$ are to be used as possible explanatory variables in a linear model, the use of principal components gives the additional advantage that the Z_j's are mutually uncorrelated. A discussion of the use of principal components in regression analysis is given in [8] along with some examples dealing with economic modelling.

The vectors $\phi_1,...,\phi_p$ form a basis for the p dimensional Euclidean space R_p. If $\psi_1,...,\psi_p$ is any other orthonormal basis for R_p, it can be shown [10 p. 400] that for any $k \leq p$,

$$E\{\sum_{i=1}^{p} | x_i - \sum_{\nu=1}^{k} Z_\nu \phi_{i\nu} |^2\} \leq E\{\sum_{i=1}^{p} | x_i - \sum_{\nu=1}^{k} U_\nu \psi_{i\nu} |^2\},$$

$$(1.1)$$

where $\phi_i = \begin{bmatrix} \phi_{i1} \\ \cdot \\ \cdot \\ \cdot \\ \phi_{ip} \end{bmatrix}$, $\psi_i = \begin{bmatrix} \psi_{i1} \\ \cdot \\ \cdot \\ \cdot \\ \psi_{ip} \end{bmatrix}$ and $U_\nu = \psi_\nu'X$. or

that the k components $Z_1,\ldots,Z_k$ give a better approximation in the sense of expected squared error to $x_1,\ldots,x_p$ than do the k coefficients of any other orthogonal basis. Moreover, when X is scaled so that $\sum_{i=1}^{p} E(x_i^2) = 1$, then,

$$-\sum_{i=1}^{k} \lambda_i \log \lambda_i \leq - \sum_{i=1}^{k} \rho_i \log \rho_i,$$

$$(1.2)$$

where $\rho_i = E(U_i^2)$. Equality holds in (1.2) only when $\sum_{i=1}^{k} \lambda_i = \sum_{i=1}^{k} \rho_i$ for every $0 < k < p$. (1.1) and (1.2) show that for any fixed k < p, the first k principal components carry at least as much of the information contained in the set $x_1,\ldots,x_p$ of variables as any k linear combinations, including any subset $x_{i_1},\ldots,x_{i_k}$.

When $x_1,\ldots,x_p$ are to be considered as possible explanatory variables in a linear model for another variable, say, y, then it is of interest to be able to characterize when the principal components $Z_1,\ldots,Z_k$ provide more information relevant to y than some other linear combinations $U_1,\ldots,U_k$.

Let y, $X' = (x_1,\ldots,x_p)$ be given random variables such that $Ey = E(x_i) = 0$, $i=1,\ldots,p$; $EXX' = \Sigma$, $EyX' = \Sigma_{12}$, $\Sigma_{12}' = \Sigma_{21}$. As before let $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_p$ and $\phi_1,\ldots,\phi_p$ be the eigenvalues and corresponding orthonormalized eigenvectors of Σ. Let e_i denote the vector

$$\begin{pmatrix} 0 \\ \cdot \\ \cdot \\ \cdot \\ 0 \\ 1 \\ 0 \\ \cdot \\ \cdot \\ \cdot \\ 0 \end{pmatrix} \text{ith row}$$

and $\hat{\beta}$ be the regression vector of y on X (i.e. the vector minimizing the expression $E(y - X'\beta)^2$.) Suppose now that y is to be predicted by a linear combination of $k < p$ random variables which are themselves linear combinations of $x_1,\ldots,x_p$. One possibility is to choose some subset $x_{i_1},\ldots,x_{i_k}$, while another is to choose the first k principal components $Z_1,\ldots,Z_k$.

Let $Z = \begin{bmatrix} z_1 \\ \cdot \\ \cdot \\ \cdot \\ z_k \end{bmatrix} = \phi^{(k)'} X$, where $\phi^{(k)} = (\phi_1, \ldots, \phi_k)$, and

let $W = \begin{bmatrix} w_1 \\ \circ \\ \cdot \\ \cdot \\ w_k \end{bmatrix} = E^{(k)'} X$, where $E^{(k)} = (e_{i_1}, \ldots, e_{i_k})$.

Let $\hat{y}_Z = Z'\hat{a}$ denote the predictor of y based on Z,

where $\hat{a} = \begin{bmatrix} \hat{a}_1 \\ \cdot \\ \cdot \\ \cdot \\ \hat{a}_k \end{bmatrix}$ is the vector minimizing $E(y - Z'a)^2$,

and let $\hat{y}_W = W'\hat{c}$ be the predictor of y based on W,
where $\hat{c}$ is the vector minimizing $E(y - W'c)^2$. Then

$$\hat{y}_Z = X'\phi^{(k)}\hat{a} = X'\phi, \quad \text{while} \quad \hat{y}_W = X'E^{(k)}\hat{c} = X'\psi.$$

Definition: Let u be a vector in R_p and define the norm $||u||$ by $||u||^2 = (u,u)$, where $(u,v) = u'\Sigma v$.

Lemma (1.1): $E(y - \hat{y}_Z)^2 < E(y - \hat{y}_W)^2$ iff $||\hat{\beta} - \phi|| < ||\hat{\beta} - \psi||$.

Proof: $E(y - \hat{y}_Z)^2 = E(y - X'\phi)^2 = E\{(y - X'\hat{\beta})$

$+ X'(\hat{\beta} - \phi)\}^2 = E(y - X'\hat{\beta})^2 = E[(\hat{\beta} - \phi)'XX'(\hat{\beta}-\phi)]$

$= E(y - X'\hat{\beta})^2 + ||\hat{\beta} - \phi||^2$. Similarly, $E(y-\hat{y}_W)^2$

$= E(y - X'\hat{\beta})^2 + ||\hat{\beta} - \psi||^2$ from which the lemma

follows.

Theorem 1.1: $E(y - \hat{y}_Z)^2 < E(y - \hat{y}_W)^2$ iff

$\Sigma_{12}(\Sigma^{-1} - \sum\limits_{i=1}^{k} \frac{1}{\lambda_i}\phi_i\phi_i')\Sigma_{21} < \Sigma_{12}(\Sigma^{-1} - \tilde{\Sigma}_k)\Sigma_{21}$, where

$$\tilde{\Sigma}_k = \begin{bmatrix} 0 & \cdots & & & & & & & & & & \cdots & 0 \\ \vdots & & & & & & & & & & & & \\ 0 & \cdots & \sigma^{i_1 i_1} & 0 & \cdots & \sigma^{i_1 i_2} & 0 & \cdots & \sigma^{i_1 i_k} & 0 & \cdots & & 0 \\ \vdots & & & & & & & & & & & & \\ 0 & \cdots & \sigma^{i_2 i_1} & 0 & \cdots & \sigma^{i_2 i_2} & 0 & \cdots & \sigma^{i_2 i_k} & 0 & \cdots & & 0 \\ \vdots & & & & & & & & & & & & \\ 0 & \cdots & \sigma^{i_k i_1} & 0 & \cdots & \sigma^{i_k i_2} & 0 & \cdots & \sigma^{i_k i_k} & 0 & \cdots & & 0 \\ \vdots & & & & & & & & & & & & \\ 0 & \cdots & & & & & & & & & & \cdots & 0 \end{bmatrix},$$

and where $\sigma^{i_j i_k}$ is the jkth element of S_k^{-1} where

$$S_k = \begin{bmatrix} \sigma_{i_1 i_1} & \cdots & \sigma_{i_1 i_k} \\ \vdots & & \\ \vdots & & \\ \sigma_{i_k i_1} & \cdots & \sigma_{i_k i_k} \end{bmatrix}.$$

Proof: $||\hat{\beta} - \phi||^2 = (\hat{\beta} - \phi, \hat{\beta} - \phi) = (\hat{\beta}, \hat{\beta}) - 2(\hat{\beta}, \phi)$

$+ (\phi, \phi)$. $(\hat{\beta}, \hat{\beta}) = \Sigma_{12} \Sigma^{-1} \Sigma_{21}$, $(\hat{\beta}, \phi) = \hat{\beta}' \Sigma \phi$

$= \hat{\beta}' \Sigma \phi^{(k)} \hat{a} = \Sigma_{12} \Sigma^{-1} \Sigma \phi^{(k)} \hat{a} = \Sigma_{12} \phi^{(k)} \hat{a}$. But,

$\hat{a} = [E(ZZ')]^{-1} E(Zy) =$

$[E(\phi^{(k)'} XX' \phi^{(k)})]^{-1} E(\phi^{(k)'} Xy)$

$= [\phi^{(k)'} \Sigma \phi^{(k)}]^{-1} \phi^{(k)'} \Sigma_{21}$

$$= \left[\begin{bmatrix} \phi_1' \\ \vdots \\ \phi_k' \end{bmatrix} (\Sigma\phi_1 \ldots \Sigma\phi_k) \right]^{-1} \phi^{(k)'} \Sigma_{21}$$

$$= \begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_k \end{bmatrix}^{-1} \phi^{(k)'} \Sigma_{21} = \Lambda^{-1} \phi^{(k)'} \Sigma_{21},$$

so that, $(\hat{\beta}, \phi) = \Sigma_{12} \phi^{(k)} \Lambda^{-1} \phi^{(k)'} \Sigma_{21}$.

$(\phi, \phi) = \phi \Sigma \phi = \hat{a}' \phi^{(k)'} \Sigma \phi^{(k)} \hat{a}$

$= \Sigma_{12} \phi^{(k)} \Lambda^{-1} \phi^{(k)'} \Sigma \phi^{(k)} \Lambda^{-1} \phi^{(k)'} \Sigma_{21}$

$= \Sigma_{12} \phi^{(k)} \Lambda^{-1} \Lambda \Lambda^{-1} \phi^{(k)'} \Sigma_{21} = \Sigma_{12} \phi^{(k)} \Lambda^{-1} \phi^{(k)'} \Sigma_{21}$.

Hence, $||\hat{\beta} - \phi||^2 = \Sigma_{12}\Sigma^{-1}\Sigma_{21} - \Sigma_{12}\phi^{(k)}\Lambda_{-1}\phi^{(k)'}\Sigma_{21}$

$$= \Sigma_{12}\Sigma^{-1}\Sigma_{21} - \Sigma_{12}(\sum_{i=1}^{k} \frac{1}{\lambda_i}\phi_i\phi_i')\Sigma_{21}$$

$$= \Sigma_{12}(\Sigma^{-1} - \sum_{i=1}^{k}\frac{1}{\lambda_i}\phi_i\phi_i')\Sigma_{21}.$$

Similarly, $||\hat{\beta} - \Psi||^2 = (\hat{\beta},\hat{\beta}) - 2(\hat{\beta},\Psi) + (\Psi,\Psi).$

$(\hat{\beta},\Psi) = \hat{\beta}'\Sigma E^{(k)}\hat{c},$ where $\hat{c} = [E(WW')]^{-1}E(Wy)$

$$= [E(E^{(k)'}XX'E^{(k)})]^{-1} E(E^{(k)'}Xy)$$

$$= [E^{(k)'}\Sigma E^{(k)}]^{-1} E^{(k)'}\Sigma_{21}$$

$$= [\begin{pmatrix} e_{i_1}' \\ \cdot \\ \cdot \\ \cdot \\ e_{i_k}' \end{pmatrix} (\Sigma e_{i_1}\ldots\Sigma e_{i_k})]^{-1} E^{(k)'}\Sigma_{21}$$

$$= \begin{pmatrix} \sigma_{i_1 i_1}\cdots\sigma_{i_1 i_k} \\ \cdot \\ \cdot \\ \cdot \\ \sigma_{i_k i_1}\cdots\sigma_{i_k i_k} \end{pmatrix}^{-1} E^{(k)'}\Sigma_{21} = S_k^{-1}E^{(k)'}\Sigma_{21}.$$

Thus, $(\hat{\beta},\Psi) = \Sigma_{12}\Sigma^{-1}\Sigma E^{(k)}S_k^{-1}E^{(k)'}\Sigma_{21}.$

$$= \Sigma_{12}E^{(k)}S_k^{-1}E^{(k)'}\Sigma_{21}, \text{ while } (\Psi,\Psi) = \Psi'\Sigma\Psi$$

$$= \Sigma_{12} E^{(k)} S_k^{-1} E^{(k)'} \Sigma E^{(k)} S_k^{-1} E^{(k)'} \Sigma_{21}$$

$$= \Sigma_{12} E^{(k)} S_k^{-1} S_k S_k^{-1} E^{(k)'} \Sigma_{21} = \Sigma_{12} E^{(k)} S_k^{-1} E^{(k)'} \Sigma_{21}.$$

Hence $||\hat{\beta} - \psi||^2 = \Sigma_{12} \Sigma^{-1} \Sigma_{21} - \Sigma_{12} E^{(k)} S_k^{-1} E^{(k)'} \Sigma_{21}$

$$= \Sigma_{12} [\Sigma^{-1} - \tilde{\Sigma}_k] \Sigma_{21},$$

and the theorem follows from lemma 1.1.

In practice if $x_1, \ldots, x_p$ are highly inter-correlated, then it might be expected that $\sum\limits_{i=1}^{k} \frac{1}{\lambda_i} \phi_i \phi_i'$ would more closely approximate Σ^{-1} then would $\tilde{\Sigma}_k$. If at the same time, y is essentially equally correlated with all or many of the x_j's, then it would seem likely that the inequality would hold. In particular, this situation might arise when $x_1, \ldots, x_p$ are elements of a time series $x(t), x(t - 1), \ldots, x(t - p + 1)$ and $y = y(t)$.

As an example, suppose $x(t)$ is a first order autoregressive process with $x(t) - .9x(t - 1) = Z(t)$ (a white noise process). Then the covariance matrix Σ of $x(t), x(t - 1), x(t - 2)$ can be taken as

$$\Sigma = \begin{bmatrix} 1.0 & .9 & .81 \\ .9 & 1.0 & .9 \\ .81 & .9 & 1.0 \end{bmatrix}, \text{ with eigenvalues } \lambda_1 = 2.74,$$

$\lambda_2 = .19$, $\lambda_3 = .07$ and eigenvectors

$$\phi_1 = \begin{bmatrix} .571 \\ .590 \\ .571 \end{bmatrix} \quad , \quad \phi_2 = \begin{bmatrix} -.707 \\ 0.0 \\ .707 \end{bmatrix} \quad , \quad \phi_3 = \begin{bmatrix} -.417 \\ .807 \\ -.417 \end{bmatrix} \quad .$$

$$\Sigma^{-1} = \begin{bmatrix} 5.232 & -4.687 & -.027 \\ -4.687 & 9.431 & -4.687 \\ -.027 & -4.687 & 5.232 \end{bmatrix} \quad . \text{ Letting } k = 1,$$

$$\frac{1}{\lambda_1} \phi_1 \phi_1' = \begin{bmatrix} .117 & .121 & .117 \\ .121 & .125 & .121 \\ .117 & .121 & .117 \end{bmatrix} , \text{ and}$$

$$\tilde{\Sigma}_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad . \text{ Letting } \Sigma_{12} = (1 \quad .99 \quad .98),$$

$$\Sigma_{12}(\Sigma^{-1} - \frac{1}{\lambda_1} \phi_1 \phi_1')\Sigma_{21} = .017 \text{ while}$$

$$\Sigma_{12}(\Sigma^{-1} - \tilde{\Sigma}_1)\Sigma_{21} = .071.$$

CHAPTER II

STOCHASTIC PROCESSES AND REPRESENTATIONS

The type of stochastic process $x(t)$ to be considered will be a family of random variables with index set T, where unless otherwise stated T will denote a countable set of the real line, so that $x(t)$ is a discrete process. The random variables $x(t)$ may be real or complex valued, but will always be assumed to satisfy the condition $E|x(t)|^2 < \infty$ for all $t \in T$. The mean value function of $x(t)$ will be denoted by $\mu(t)$, and unless otherwise stated it will be assumed that $\mu(t) \equiv 0$. The covariance function will be denoted by $r(s,t)$ $= E\{x(s) \overline{x(t)}\}$ where the bar denotes complex conjugate. Of particular importance is the case where $x(t)$ is a second order stationary process. In that case the covariance function has the form

$$E\{x(s) \overline{x(s-t)}\} = r(t),$$

and $r(t)$ has the spectral representation

14

$$r(t) = \int_{-\pi}^{\pi} e^{it\omega} \, dF(\omega),$$

where $F(\omega)$, the spectral distribution function of $x(t)$, is real valued, nondecreasing, bounded, and can be chosen so that $F(-\infty) = 0$, $F(+\infty) = r(0)$, and $F(\omega)$ is continuous to the right. If $F(\omega)$ is absolutely continuous then $f(\omega) = F'(\omega)$ is the spectral density function of $x(t)$. In addition, if $x(t)$ is second order stationary, it has the spectral representation

$$x(t) = \int_{-\pi}^{\pi} e^{it\omega} \, d\xi(\omega),$$

where $\xi(\omega)$ is a process of orthogonal increments which can be chosen so that $E\,\xi(\omega) = 0$, $E|\xi(\omega)|^2 = F(\omega)$, and $E|d\xi(\omega)|^2 = dF(\omega)$. ($\xi(\omega)$ is called the spectral process associated with $x(t)$.) From this spectral representation the process $x(t)$ can be thought of as being built up of mutually orthogonal elementary harmonic oscillations $e^{it\omega} \, d\xi(\omega)$.

Let $x(t)$ be a second order stationary process and consider a process $y(t)$ formed from $x(t)$ by a linear transformation of the type

$$y(t) = \sum_{\nu=1}^{n} b_\nu \, x(t - t_\nu). \qquad (2.1)$$

Then

$$y(t) = \sum_{\nu=1}^{n} b_\nu \int_{-\infty}^{\infty} e^{i(t-t\nu)\omega} d\xi_x(\omega)$$

$$= \int_{-\infty}^{\infty} e^{it\omega} \left(\sum_{\nu=1}^{n} b_\nu e^{-it_\nu\omega} \right) d\,\xi_x(\omega).$$

Thus, $y(t)$ is a second order stationary process with elementary harmonic oscillations $h(\omega) \, d\,\xi_x(\omega)$, where $h(\omega)$ $= \sum_{\nu=1}^{n} b_\nu e^{-it_\nu\omega}$. The covariance function $r_y(t)$ of $y(t)$ has the representation

$$r_y(t) = \int_{-\infty}^{\infty} e^{it\omega} \, |h(\omega)|^2 \, d\, F_x(\omega).$$

If $x(t)$ has spectral density function $f_x(\omega)$ then $y(t)$ has spectral density function $f_y(\omega) = |h(\omega)|^2 \, f_x(\omega)$. Thus for a given frequency ω, $f_x(\omega)$ is multiplied by the amount $|h(\omega)|^2$ to give $f_y(\omega)$. This may have the effect of greatly repressing certain frequencies while enhancing others. For this reason the linear operation (2.1) may be thought of as a filter applied to the process

$x(t)$. $|h(\omega)|^{2}$ is called the transfer function of the filter and $h(\omega)$ the gain of the filter.

Karhunen - Loève Representations

Theorem [see 7 p. 478]

Let $x(t)$ be a stochastic process continuous in quadratic mean and defined for $t \in T$ a closed interval. Then $x(t)$ has an orthogonal representation of the form

$$x(t) = \sum_{n=1}^{\infty} \psi_n(t) \, Z_n,$$

where $\psi_n(t)$ is an eigenfunction of the covariance function $R(t,t') = E\{x(t) \, x(t')\}$. That is

$$\int_T R(t,t') \, \psi_n(t') \, dt' = \lambda_n \, \psi_n(t),$$

and

$\int_T \psi_m(t) \, \overline{\psi_n}(t) \, dt = \delta_{mn}$. The random variables Z_n are given by

$$Z_n = \int_T x(t) \, \overline{\psi_n}(t) \, dt,$$

and the Z_n's have the property that $E(Z_m \, \overline{Z}_n) = \lambda_n \, \delta_{mn}$.

When $x(t)$ is a finite discrete process, T can be taken as the set $T = \{1, 2, \ldots, N\}$, and $R(t,t')$ becomes the

N X N matrix

$$R^{(N)} = \begin{bmatrix} R(1,1) & \dots & R(1,N) \\ R(2,1) & \dots & R(2,N) \\ \cdot \\ \cdot \\ \cdot \\ R(N,1) & \dots & R(N,N) \end{bmatrix} ,$$

with eigenvalues $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_N$, and corresponding orthonormal set of eigenvectors $\psi_1, \dots, \psi_N$, where

$$\psi_k = \begin{bmatrix} \psi_k(1) \\ \cdot \\ \cdot \\ \cdot \\ \psi_k(N) \end{bmatrix} .$$ The representation for x(t) then

becomes

$$x(t) = \sum_{n=1}^{N} \psi_n(t) \ Z_n,$$

where

$$Z_n = \sum_{t=1}^{N} x(t) \overline{\psi_n}(t).$$

Thus in this case Z_n is the nth principal component of the random variables $x(1), \dots, x(N)$.

Principal Component Processes

Let $x(t)$, $t = 1,2,\dots$ be a discrete stochastic process.

Definition: $x(t)$ will be called covariance stationary if $E\{(x(t) - \mu(t)) \ (\overline{x(s)} - \overline{\mu(s)})\} = r(t - s)$, where

$\mu(t) = E\ x(t)$. ($\mu(t)$ is not necessarily constant.) For any fixed positive integer n let $R^{(n)}$ be the covariance matrix

$$R^{(n)} = \begin{bmatrix} r(0) & r(1) & . & . & . & . & r(n) \\ r(-1) & r(0) & . & . & . & . & r(n-1) \\ . \\ . \\ . \\ r(-n) & r(-n+1) & . & . & r(0) \end{bmatrix} ,$$

$\lambda_0^{(n)} \geq \lambda_1^{(n)} \geq \lambda_2^{(n)} \geq \ldots \geq \lambda_n^{(n)}$ be the eigenvalues of $R^{(n)}$, and $\phi_0^{(n)}, \phi_1^{(n)}, \ldots, \phi_n^{(n)}$ be the corresponding set of orthonormalized eigenvectors.

Definition: The processes $Z_j^{(n)}(t)$; $j = 0,1,\ldots,n$ where

$$Z_j^{(n)}(t) = \sum_{\nu=0}^{n} \phi_j^{(n)}(\nu)\ x(t-\nu)$$

will be called the principal component processes of $x(t)$.

From this definition and the usual properties of principal components it follows that:

(1) $Z_0(t)$ is that linear combination

$\sum_{i=0}^{n} a_{i0}\ x(t-1)$ with maximum variance sub-

ject to the restriction that $\sum_{i=1}^{n} a_{i0}^2 = 1$,

$Z_1^{(n)}(t)$ is the linear combination

$\sum_{i=1}^{n} a_{i1}\ x(t-1)$ with maximum variance

subject to the restrictions that $\sum\limits_{i=1}^{n} a_{i1}^2 = 1$

and $\sum\limits_{i=1}^{n} a_{i0} \overline{a_{i1}} = 0$, and so forth for

$Z_j^{(n)}(t)$, $j = 0, \ldots, n$.

(ii) $E\{Z_i^{(n)}(t) \overline{Z_j^{(n)}(t)}\} = \lambda_i \delta_{ij}$.

(iii) $x(t)$ has the representation,

$$x(t) = \sum_{\nu=0}^{n} \overline{\phi_\nu(0)}\, Z_\nu^{(n)}(t).$$

This representation differs in general from
the Karhunen-Loève representation or the
spectral representation in that the index
t is contained in the random element. Thus
the representation is in terms of constants
and random processes rather than random
variables and deterministic functions of t.

(iv) $r(0) = \sum\limits_{\nu=0}^{n} |\phi_\nu(0)|^2 \lambda_\nu$.

An interpretation of these processes can be seen
more readily in the case that $x(t)$ is a second order

stationary process. Let $x(t)$ be a discrete stationary process with spectral process $\xi_x(\omega)$, and assume more-over, that $x(t)$ has a spectral density function $f_x(\omega)$. Then

$$z_j^{(n)}(t) = \sum_{\nu=0}^{n} \phi_j^{(n)}(\nu)\, x(t-\nu), \quad j = 0,\ldots,n$$

is a linear operation on $x(t)$ and hence, it follows that,

(v) $z_j^{(n)}(t)$ is a stationary process,

(vi) $z_j^{(n)}(t)$ is built up of elementary harmonic

oscillations of the form $e^{it\omega}\, g_\nu^{(n)}(\omega)\, d\xi_x(\omega)$

where $g_\nu^{(n)}(\omega) = \sum_{k=0}^{\cdot} e^{-ik\omega}\, \phi_\nu^{(n)}(k)$,

(vii) $z_j^{(n)}(t)$ has spectral density function

$$f_{z_j^{(n)}}(\omega) = |g_j^{(n)}(\omega)|^2\, f_x(\omega).$$

From (iii), (vi), (vii) it follows that the spectral representations for $x(t)$ and $r(0)$ may be written in the

forms

$$x(t) = \int_{-\Pi}^{\Pi} e^{it\omega} d\xi_x(\omega)$$

$$= \int_{-\Pi}^{\Pi} e^{it\omega} \overline{(\phi_0(0)} g_0(\omega) d\xi_x(\omega) + \ldots$$

$$+ \overline{\phi_n(0)} g_n(\omega) d\xi_x(\omega)),$$

and

$$r(0) = \int_{-\Pi}^{\Pi} f_x(\omega) d\omega$$

$$= \int_{-\Pi}^{\Pi} (|\phi_0(0)|^2 |g_0(\omega)|^2 f_x(\omega) + \ldots$$

$$+ |\phi_n(0)|^2 |g_n(\omega)|^2 f_x(\omega)) d\omega$$

which gives a decomposition for the elementary harmonic

oscillations $e^{it\omega} d\xi_x(\omega)$ and the spectral density func-

tion $f_x(\omega)$. From the properties (i) and (ii) it follows

that

$$\int_{-\Pi}^{\Pi} |g_0(\omega)|^2 f_x(\omega) d\omega = \omega_0 \geq \int_{-\Pi}^{\Pi} |h(\omega)|^2 f_x(\omega) d\omega,$$

for any $h(\omega)$ satisfying

$$h(\omega) = \sum_{\nu=0}^{n} b_\nu e^{-i\nu\omega}, \quad \sum_{\nu=0}^{n} b_\nu^2 = 1.$$

Thus, the filter function $g_0(\omega)$ passes the most import-
ant portion of the spectrum of $x(t)$, while $g_1(\omega),\ldots,g_n(\omega)$
pass increasingly smaller portions of $f_x(\omega)$. The exact
properties of these filter functions are determined by
the filter coefficients

$$\phi_\nu(k), \quad k = 0,\ldots,n; \quad \nu=0,\ldots,n$$

which in turn are determined by the covariance structure
of the process $x(t)$. These properties are further illus-
trated by the following special case.

x(t) A Periodic Process

Let $x(t)$ be a stationary process such that
$x(t + N) = x(t)$ for some integer N. This process can be
thought of as containing only a finite number of elemen-
tary harmonic oscillations and, hence, has the represen-
tation

$$x(t) = \sum_{j=0}^{N-1} e^{i\left(\frac{2\pi j}{N}\right)t} x_j, \tag{2.2}$$

where the

$$x_j = \frac{1}{N} \sum_{k=0}^{N-1} e^{-i\left(\frac{2\pi j}{N}\right)} x(k), \quad j = 0,\ldots,n-1 \text{ are}$$

mutually orthogonal random variables. (2.2) can be put in the more usual form by letting

$$\omega_j = \begin{cases} \dfrac{2\pi j}{N} , & j \leq \dfrac{N}{2} \\[2mm] \dfrac{2\pi j}{N} - 2\pi, & j > \dfrac{N}{2} \end{cases} .$$

Then (2.2) becomes

$$x(t) = \sum_{j=0}^{N-1} e^{i\omega_j t} x_j, \qquad -\pi \leq \omega_j \leq \pi.$$

From the definition of x(t), $E\, x(t)\, \overline{x(t - (N - j))}$
$= E\, x(t)\, \overline{x(t + j)}$ or $r(n - j) = r(-j)$. Hence, the covariance matrix

$$R^{(N-1)} = \begin{pmatrix} r(0) & r(1) & \ldots & r(N-1) \\ r(-1) & r(0) & \ldots & r(N-2) \\ \cdot & & & \\ \cdot & & & \\ \cdot & & & \\ r(-N+1) & r(-N+2) & \ldots & r(0) \end{pmatrix}$$

$$= \begin{pmatrix} r(0) & r(1) & \ldots & r(N-1) \\ r(N-1) & r(0) & \ldots & r(N-2) \\ \cdot & & & \\ \cdot & & & \\ \cdot & & & \\ r(1) & r(2) & \ldots & r(0) \end{pmatrix}$$

is a circular matrix. As such it has eigenvalues given by

$$\lambda_j = \sum_{k=0}^{N-1} \rho_j^{\,k}\, r(k), \text{ where } \rho_j \text{ is an Nth root of}$$

unity. The corresponding eigenvector is given by

$$\phi_j^{(n)} = \begin{bmatrix} 1 \\ \rho_j \\ \rho_j^2 \\ \cdot \\ \cdot \\ \cdot \\ \rho_j^{N-1} \end{bmatrix} \frac{1}{\sqrt{N}}$$

Putting $\rho_j = e^{\frac{i2\pi j}{N}}$ gives

$$\phi_j^{(n)} = \begin{bmatrix} 1 \\ e^{\frac{i2\pi j}{N}} \\ e^{\frac{i2\pi 2j}{N}} \\ \cdot \\ \cdot \\ \cdot \\ e^{\frac{i2\pi(N-1)j}{N}} \end{bmatrix} \frac{1}{\sqrt{N}} \qquad (2.3)$$

The principal component process $Z_j^{(n)}(t)$ then has the form

$$Z_j^{(n)}(t) = \Sigma e^{\frac{i2\pi j\nu}{N}} x(t-\nu) \frac{1}{\sqrt{N}}$$

$$= \frac{1}{\sqrt{N}} \sum_{\nu=0}^{N-1} e^{\frac{i2\pi j\nu}{N}} \sum_{k=0}^{N-1} e^{\frac{i2\pi k(t-\nu)}{N}} X_k$$

$$= \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} \sum_{\nu=0}^{N-1} e^{\frac{i2\pi(j-k)\nu}{N}} e^{\frac{i2\pi kt}{N}} X_k$$

$$= \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} N\delta_{jk} e^{i\frac{2\pi k}{N}t} \qquad x_k = \sqrt{N} e^{i\frac{2\pi j}{N}t} x_j,$$

so that $Z_j(t)$ consists only of the elementary harmonic oscillation at frequency $\frac{2\pi j}{N}$. This may also be seen from the filter function corresponding to $Z_j(t)$, which is

$$g_j(\omega) = \frac{1}{\sqrt{N}} \sum_{\nu=0}^{N-1} e^{i\frac{2\pi j}{N}} e^{-i\nu\omega}$$

$$= \frac{1}{\sqrt{N}} \sum_{\nu=0}^{N-1} e^{i(\frac{2\pi j}{N} - \omega)\nu}$$

But $x(t)$ only contains oscillations at frequencies

$\omega_k = \frac{2\pi k}{N}$, $k=0,\ldots,N-1$ and $g_j(\omega_k) = \sqrt{N}\,\delta_{jk}$.

The representation (iii) in this case yields

$$x(t) = \sum_{j=0}^{N-1} \frac{1}{\sqrt{N}} (\sqrt{N} e^{i\frac{2\pi j}{N}t} x_j)$$

$$= \sum_{j=0}^{N-1} e^{i\frac{2\pi j}{N}t} x_j$$

which is identical with the spectral representation (2.2). On the other hand, since $x(t + N) = x(t)$, the process $x(t)$ consists of only a finite number of distinct random variables. Thus, considering only the distinct $x(t)$'s, the

representation (2.2) is identical with the Karhunen representation with $Z_j = X_j$.

Since an arbitrary process may be approximated by a periodic one by allowing the period to tend to infinity, it seems reasonable to inquire whether some of the properties exhibited by the principal component processes for the periodic case might not hold in some limiting sense for arbitrary processes. For instance, the form of the eigenvalues of the circular matrix

$$\lambda_j^{(N-1)} = \sum_{k=0}^{N-1} (e^{i\frac{2\pi j}{N}})^k r(k)$$

might suggest that the set $\{\lambda_v^{(n)}\}_{v=0}^{n}$ of eigenvalues of the covariance matrix of a stationary process might in some way approximate its spectral density function.

These ideas are made more precise using the theory of Toeplitz forms developed by Grenander and Szegö [5]. Definition: Let $f(x)$ be a real valued function defined on $[-\Pi, \Pi]$ such that $\int_{-\Pi}^{\Pi} |f(x)| \, dx$ exists, and let $c^{(n)}$ be the matrix $(c_{v-\mu})$; $\mu, v = 0,\ldots,n$, where

$$c_k = \frac{1}{2\Pi} \int_{-\Pi}^{\Pi} e^{-ikx} f(x) dx.$$

Then $c^{(n)}$ will be called the Toeplitz matrix of order n associated with the function $f(x)$. Similarly, if $F(x)$ is monotone nondecreasing in $[-\Pi,\Pi]$, the Toeplitz matrix associated with $F(x)$ is $c^{(n)} = (c_{\nu-\mu})$ where

$$c_k = \frac{1}{2\Pi} \int_{-\Pi}^{\Pi} e^{-ikx} \, dF(x).$$

The quadratic form $X \, c^{(n)} \, X^*$ is called the Toeplitz form associated with $f(x)$ or $F(x)$ where $X = (x_0, x_1, \ldots, x_n)$ is an arbitrary complex vector.

Proposition: Let $x(t)$ be a second order stationary stochastic process and $R^{(n)}$ the covariance matrix

$$R^{(n)} = \begin{bmatrix} r(0) & \ldots & r(n) \\ \vdots & & \\ \vdots & & \\ r(-n) & \ldots & r(0) \end{bmatrix}$$

Then $R^{(n)}$ is the Toeplitx matrix associated with the spectral distribution $2\Pi \, F_x(\omega)$ (or spectral density function $2\Pi \, f(\omega)$ if it exists).

Proof: Bochner's theorem for the discrete case gives

$$r(n) = \int_{-\Pi}^{\Pi} e^{in\omega} \, dF_x(\omega) \text{ so that}$$

$$r(-n) = \frac{1}{2\Pi} \int_{-\Pi}^{\Pi} e^{-in\omega} \, d[2\Pi \, F_x(\omega)].$$

<u>Definition:</u> For each n consider the two sets of real

numbers $\{a_\nu^{(n)}\}_{\nu=0}^n$ and $\{b_\nu^{(n)}\}_{\nu=0}^1$, where there exists a

constant K independent of n and ν, such that $|a_\nu^{(n)}| < K$,

$|b_\nu^{(n)}| < K$. $\{a_\nu^{(n)}\}$ and $\{b_\nu^{(n)}\}$, $n \to \infty$ are said to be

equally distributed in $[-K, K]$ if

$$\lim_{n \to \infty} \sum_{\nu=0}^n \frac{[F(a_\nu^{(n)}) - F(b_\nu^{(n)})]}{n+1} = 0,$$

where $F(t)$ is an arbitrary continuous function on $[-K, K]$.

<u>Theorem:</u> Let $f(x)$ be a real valued function such that

$\int_{-\Pi}^{\Pi} |f(x)| dx < \infty$, and let m and M denote the essential

lower and upper bounds (assumed finite) respectively of

$f(x)$. If $F(\lambda)$ is any continuous function on $[m, M]$, then

$$\lim_{n \to \infty} \frac{F(\lambda_0^{(n)}) + \ldots + F(\lambda_n^{(n)})}{n+1} = \frac{1}{2\Pi} \int_{-\Pi}^{\Pi} F(f(x)) dx,$$

$$(2.4)$$

where $\lambda_0^{(n)} \geq \ldots \geq \lambda_n^{(n)}$ are the eigenvalues of the

matrix $c^{(n)}$ associated with $f(x)$. For proof see [5].
By considering the integral on the right hand side of
(2.4) as a limit of approximating sums, the theorem shows
that the sets $\{\lambda_\nu^{(n)}\}_{\nu=0}^n$ and $\{f(\Pi- \frac{2(\nu+1)\Pi}{n+2})\}_{\nu=0}^n$ are
equally distributed as $n \to \infty$, and by proper choice of
$F(\lambda)$ it follows that

$$\lim_{n\to\infty} \lambda_0^{(n)} = M. \qquad (2.5)$$

These results show that when $x(t)$ is a second order
stationary process with spectral density function $f_x(\omega)$,
the eigenvalues $\lambda_0^{(n)},\ldots,\lambda_n^{(n)}$ of $R^{(n)}$ tend to approxi-
mate $2\Pi f_x(\omega)$ at an equally spaced set of points in
$[-\Pi,\Pi]$, and in particular,

$$\lim_{n\to\infty} \lambda_0^{(n)} = \max_{-\Pi \le \omega \le \Pi} 2\Pi f_x(\omega)$$

Pertaining to the asymptotic behavior of eigenvectors, it
is shown in [5] that a Toeplitz matrix can be approxi-
mated according to a certain norm by a hermitian matrix
(with known eigenvalues and eigenvectors) in such a way
that the distributions of eigenvalues of the two matrices
are asymptotically equal. The mode of convergence, how-
ever, does not seem to be strong enough to guarantee the
convergence of a transfer function given by an eigenvector

of a Toeplitz matrix to the same limit as that of the transfer function given by the corresponding eigenvector of the approximating matrix. In particular, it is shown that an arbitrary Toeplitz matrix can be approximated by a matrix with eigenvectors given by (2.3). This might suggest that the principal component processes might, in general, tend to give a frequency decomposition as in the case of the periodic process. That this is not always the case can be seen by considering a process $x(t)$ of the opposite extreme.

x(t) A Discrete White Noise Process

Let $x(t)$ be such that $E\,x(t) = 0$, and $E\{x(t)\ \overline{x(t')}\} = \sigma^2\ \delta_{tt'}$. The spectral density function of this process is

$$f_x(\lambda) = \frac{\sigma^2}{2\pi}\ , \ -\pi \le \lambda \le \pi,$$

and the covariance matrix

$$R^{(n)} = \begin{bmatrix} \sigma^2 & & & 0 \\ & \cdot & & \\ & & \cdot & \\ & & & \cdot \\ 0 & & & \sigma^2 \end{bmatrix},$$

with eigenvalues $\lambda_0^{(n)} = \ldots = \lambda_n^{(n)} = \sigma^2$ and corresponding eigenvectors $\phi_0^{(n)}, \ldots, \phi_n^{(n)}$ which may be taken as

$$\begin{bmatrix} 1 \\ 0 \\ \cdot \\ \cdot \\ \cdot \\ 0 \end{bmatrix} , \begin{bmatrix} 0 \\ 1 \\ \cdot \\ \cdot \\ \cdot \\ 0 \end{bmatrix} , \ldots, \begin{bmatrix} 0 \\ \cdot \\ \cdot \\ \cdot \\ 0 \\ 1 \end{bmatrix} . \quad \text{The transfer function}$$

given by $\phi_k^{(n)}$ is hence,

$$|g_k^{(n)}(\lambda)|^2 = |e^{-ik\lambda}|^2 = 1,$$

so that each transfer function passes each frequency with
equal weighting. Thus, while the eigenvectors in this
case are not uniquely determined (and hence neither are
the principal component processes) for at least one choice
of eigenvectors the principal component processes do not
provide a frequency decomposition of the original process.

It will be seen in what follows that for many
other common processes certain of the associated principal
component processes tend to give meaningful frequency de-
compositions.

x(t) A First Order Moving Average Process

Let $Z(t)$ be a white noise process satisfying
$E\ Z(t) = 0$, $E\{Z(t), \overline{Z(t')}\} = \delta_{tt'}$, and let
$x(t) = Z(t) + \rho Z(t-1)$, where $0 < |\rho| < 1$.

Then

$$R^{(n)} = \begin{bmatrix} 1+\rho^2 & \rho & 0 & . & . & . & . & 0 \\ \rho & 1+\rho^2 & \rho & 0 & . & . & . & 0 \\ 0 & & & & & & & \\ . & & & & & & & \\ . & & & & & & & \\ . & & & & & & & \\ 0 & . & . & . & . & 0 & \rho & 1+\rho^2 \end{bmatrix} .$$

The spectral density function $f_x(\omega)$ of $x(t)$ is

$$r_x(\omega) = \frac{1}{2\Pi} \mid 1 + \rho^2 + \rho e^{-i\omega} \mid$$

$$= \frac{1}{2\Pi} (1 + \rho^2 + 2\rho\cos\omega). \qquad (2.6)$$

From the equation of "mechanical quadrature" (see [5])

$$\frac{1}{2\Pi} \int_{-\Pi}^{\Pi} \phi(x) \, dx = \frac{1}{2N} \sum_{\nu=-N+1}^{N} \phi(\frac{\nu\Pi}{N}),$$

which holds when $\phi(x)$ is a trigonometric polynomial whose degree does not exceed $2N-1$, it can be shown that the eigenvalues of $R^{(n)}$ are given by

$$\lambda_\nu^{(n)} = (1 + \rho^2) + 2\rho\cos\frac{(\nu+1)}{n+2} \qquad (2.7)$$

with corresponding eigenvectors,

$$\hat{\varphi}_\nu(n) = \begin{bmatrix} \varphi^{(n)}_{\nu 0} \\ \cdot \\ \cdot \\ \cdot \\ \phi^{(n)}_{\nu n} \end{bmatrix} \text{, where}$$

$$\phi_{\nu k}(n) = \sqrt{\frac{2}{n+2}} \ \sin \ \frac{(k+1)(\nu+1)}{n+2} \text{ , } \nu=0,\ldots,n.$$

$$(2.8)$$

When $\rho > 0$, $\lambda_0(n) > \lambda_1(n) > \ldots > \lambda_n(n)$ and $\lambda_0(n) \xrightarrow{n\to\infty} (1+\rho)^2$

$= f_x(0) = \max_\omega f_x(\omega)$. If $\rho < 0$, the subscripts on the $\lambda^{(n)}$'s must be reversed in order to have the proper ordering. After relabeling

$$\lambda_0(n) = (1 + \rho^2) + 2\rho\cos \ \frac{(n+1)\Pi}{n+2} \xrightarrow{n\to\infty} (1 - \rho)^2$$

$$= f_x(\pm\Pi) = \max_\omega f_x(\omega).$$

For the case $n = 1$, $\rho > 0$,

$$\lambda_0(1) = (1+\rho^2) + 2\rho\cos\Pi/3 > \lambda_1(1) = (1+\rho^2) + 2\rho\cos\tfrac{2}{3}\Pi.$$

$$\phi_0^{(1)} = \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix} \text{ , } \qquad \phi_1^{(1)} = \begin{bmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \end{bmatrix} \text{ , }$$

$$Z_0^{(1)}(t) = \frac{1}{\sqrt{2}} \ x(t) + \frac{1}{\sqrt{2}} \ x(t-1),$$

$$Z_1^{(1)}(t) = \frac{1}{\sqrt{2}} \; x(t) - \frac{1}{\sqrt{2}} \; x(t-1).$$

The transfer functions corresponding to $Z_0^{(1)}(t)$ and $Z_1^{(1)}(t)$ are

$$\left| g_{Z_0}(1)^{(\omega)} \right|^2 = \left| \frac{1}{\sqrt{2}} + \frac{1}{\sqrt{2}} \; e^{-i\omega} \right|^2 = 1 + \cos\omega \text{ and}$$

$$\left| g_{Z_1}(1)^{(\omega)} \right|^2 = \left| \frac{1}{\sqrt{2}} - \frac{1}{\sqrt{2}} \; e^{-i\omega} \right|^2 = 1 - \cos\omega \text{ as shown}$$
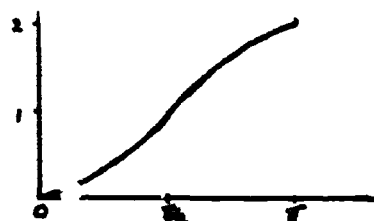
in figures 1 and 2

Figure 1

Figure 2

Similarly for any finite n, the eigenvector corresponding to the largest eigenvalue yields a principal component process whose transfer function passes th largest portion of $f_x(\omega)$.

Proposition: $\lim_{n\to\infty} f_{Z_0}(n)^{(\omega)} = 0$ for any fixed $\omega \neq 0$, $\rho > 0$

<u>Proof</u>: $g_{Z_0}(n)^{(\omega)} = \sum_{k=0}^{n} \sqrt{\frac{2}{n+2}} \left(\sin \frac{(k+1)\Pi}{n+2} \right) e^{-ik\omega}$

$$= \frac{1}{2i} \sqrt{\frac{2}{n+2}} \sum_{k=0}^{n} \left[e^{\frac{i(k+1)\Pi}{n+2}} - e^{\frac{i(k+1)\Pi}{n+2}} \right] e^{-ik\omega}$$

$$= \frac{1}{2i} \sqrt{\frac{2}{n+2}} \left\{ e^{i\frac{\Pi}{n+2}} \sum_{k=0}^{n} e^{i\left(\frac{\Pi}{n+2} - \omega\right)k} \right.$$

$$\left. -e^{-i\frac{\Pi}{n+2}} \sum_{k=0}^{n} e^{-i\left(\frac{\Pi}{n+2} + \omega\right)k} \right\}$$

$$= \frac{1}{2i} \sqrt{\frac{2}{n+2}} \left\{ e^{i\frac{\Pi}{n+2}} \frac{e^{i\left(\frac{\Pi}{n+2} - \omega\right)(n+1)} - 1}{e^{i\frac{\Pi}{n+2} - \omega} - 1} \right.$$

$$\left. - e^{\frac{-i}{n+2}} \frac{e^{-i\left(\frac{\Pi}{n+2} + \omega\right)(n+1)} - 1}{e^{-i\frac{\Pi}{n+2} + \omega} - 1} \right\}$$

$$= \frac{1}{2\Pi i} \sqrt{\frac{2}{n+2}} \quad \frac{N_{1n}(\omega)}{D_{1n}(\omega)} - \frac{N_{2n}(\omega)}{D_{2n}(\omega)} \quad . \text{ For } \omega = \omega_0 \neq 0,$$

$D_{1n}(\omega)$ and $D_{2n}(\omega) \rightarrow e^{-i\omega} - 1 \neq 0$, while $|N_{2n}(\omega)| \leq 2$ and $|N_{1n}(\omega)| \leq 2$. Thus as $n \rightarrow \infty$, $g_{Z_0}(n)^{(\omega)} \rightarrow 0$ and $f_{Z_0}(n)^{(\omega)}$

$|g_{z_0}(n)(\omega)|^2 f_x(\omega) \to 0$. Hence, the transfer function

$|g_{z_0}(n)(\omega)|^2$ acts like a δ function, $\delta(\omega)$, as $n \to \infty$. A

similar argument shows that when $\rho < 0$,

$$|g_{z_0}(n)(\omega)|^2 \to \delta(|\omega| - \Pi) \text{ as } n \to \infty.$$

x(t) A First Order Autoregressive Process

Let $Z(t)$ be a white noise process as above and

$$x(t) - \rho x(t-1) = Z(t). \qquad (2.8)$$

(2.8) may also be written

$$(1 - \rho B)x(t) = Z(t),$$

where B represents the "backwards lag operator"

(i.e. $Bx(t) = x(t-1)$), or

$$x(t) = \frac{1}{1-\rho B} \quad Z(t). \qquad (2.9)$$

If $x(t)$ is to be a stationary process then $|\rho| < 1$, in which case (2.9) may formally be written as

$$x(t) = \sum_{\nu=0}^{\infty} \rho^\nu B^\nu \quad Z(t), \text{ or}$$

$$x(t) = \sum_{\nu=0}^{\infty} \rho^\nu Z(t-\nu). \qquad (2.10)$$

Using (2.10), $R^{(n)}$ can be seen to have the form

$$R(n) = \sigma^2 \begin{bmatrix} 1 & \rho & \rho^2 & . & . & . & \rho^n \\ \rho & 1 & & . & . & . & \rho^{n-1} \\ . & & & & & & \\ . & & & & & & \\ . & & & & & & \\ \rho^n & \rho^{n-1} & . & . & . & & \rho^1 \end{bmatrix} , \qquad (2.11)$$

where $\sigma^2 = \sum_{\nu=0}^{\infty} \rho^{2\nu}$. $x(t)$ has spectral density function

$$f(\omega) = \frac{1}{2\pi} \left| \frac{1}{1-e^{i\omega}} \right|^2 = \frac{1}{2\pi} (1 + \rho^2 - 2\rho\cos\omega)^{-1} \qquad (2.12)$$

which assumes its maximum at $\omega = 0$ when $\rho > 0$. While explicit expressions for the eigenvalues and eigenvectors of $R^{(n)}$ are unfeasible, something can be said about the asymptotic properties of the largest eigenvalue and corresponding eigenvector.

Definition: A matrix $A = (a_{ij})$ is called positive if $a_{ij} > 0$ for all i,j.

Then when $\rho > 0$, $R^{(n)}$ is a positive matrix. This will be assumed to be the case in what follows.

Theorem (Perron): [see 2 p. 278] If A is a positive matrix, there is a unique eigenvalue $\lambda(A)$ which has greatest absolute value, is positive, and simple. Its associated eigenvector can be taken to be positive (have all positive elements).

__Theorem (2.0):__ [2] $\quad v = \lim\limits_{k \to \infty} \dfrac{A^k c}{[\lambda(A)]^k}$ exists and is an

eigenvector of A associated with $\lambda(A)$ for any positive

vector c. v is unique up to a scalar multiple depend-

ing on the choice of c, but otherwise, independent of

the choice of c.

__Lemma (2.1):__ Let $\phi_0 = \begin{bmatrix} \phi_{00} \\ \vdots \\ \phi_{0n} \end{bmatrix}$ be the positive eigen-

vector of $R^{(n)}$ corresponding to the largest eigenvalue.

Then $0 < \phi_{00} \le \phi_{01} \le \cdots \le \phi_{0\frac{n-1}{2}} = \phi_{0\frac{n+1}{2}} \ge \cdots \ge \phi_{0n}$

$= \phi_{00}$ if n is odd.

$0 < \phi_{00} \le \phi_{01} \le \cdots \le \phi_{0\frac{n}{2}-1} \le \phi_{0\frac{n}{2}} \ge \phi_{0\frac{n}{2}+1} = \phi_{0\frac{n}{2}-1}$

$\ge \cdots \ge \phi_{0n} = \phi_{00}$ if n is even.

__Proof:__ Only the case n even will be treated explicitly.

Positiveness can be assumed as a result of Perron's

theorem. From theorem (2.0), if c is chosen to be the

vector of 1's, then for any r

$$(R^{(n)})^r c = \begin{bmatrix} S_0{}^{(r)} \\ S_1{}^{(r)} \\ \cdot \\ \cdot \\ \cdot \\ S_n{}^{(r)} \end{bmatrix} ,$$

where $S_j{}^{(r)}$ is the jth row sum of $(R^{(n)})^r$ or $S_j{}^{(r)}$

$= \sum\limits_{v=0}^{n} b_{jv}{}^{(n)} S_v{}^{(r-1)}$, where $b_{jv}{}^{(n)}$ is the element in the

jth row, vth column of $R^{(n)}$. The equality of $S_j{}^{(r)}$

and $S_{n-j}{}^{(r)}$ is obvious from the form of $R^{(n)}$ and c. For

$r = 1$,

$S_0{}^{(1)} = 1 + \rho + \rho^2 + \ldots + \rho^n < S_1{}^{(1)} = 1 + 2\rho + \rho^2$

$+ \ldots + \rho^{n-1} < \ldots < S_{\frac{n}{2}}{}^{(1)} = 1 + 2\rho + 2\rho^2 + \ldots + 2\rho^{\frac{n}{2}}$

$> \ldots > S_n{}^{(1)} = S_0{}^{(1)}$. For $r = 2$,

$S_0{}^{(2)} = S_0{}^{(1)} + \rho S_1{}^{(1)} + \rho^2 S_2{}^{(1)} + \ldots + \rho^n S_n{}^{(1)}$ while

$S_1{}^{(2)} = \rho S_0{}^{(1)} + S_1{}^{(1)} + \rho S_2{}^{(1)} + \ldots + \rho^{n-1} S_n{}^{(1)}$, so

$S_1{}^{(2)} - S_0{}^{(2)} = (1-\rho)(S_1{}^{(1)} - S_0{}^{(1)}) + (\rho - \rho^2)S_2{}^{(1)} + \ldots$

$+ (\rho^{n-1} - \rho^n)S_n{}^{(1)} > 0$, and by the same type of argument

it can be shown that $S_0{}^{(2)} < S_1{}^{(2)} < \ldots < S_{\frac{n}{2}}{}^{(2)} > \ldots$

$> S_n^{(2)}$ by using only the elements of $R^{(n)}$ and the ordering of the $S_j^{(1)}$'s. Hence, by induction, the same ordering holds for each power r and the lemma follows by taking the limit.

Lemma 2.2: Under the hypothesis of lemma 1, $\phi_{0 \frac{n}{2}}^{(n)}$

$= \max_j \phi_{0j}^{(n)} \to 0$ as $n \to \infty$.

Proof:

$$R^{(n)} \phi_0^{(n)} = \sigma^2 \begin{bmatrix} 1 & \rho & \rho^2 & \cdots & \cdots & \rho^n \\ \rho & 1 & \rho & \cdots & \cdots & \rho^{n-1} \\ \cdot & & & & & \\ \cdot & & & & & \\ \cdot & \rho^{\frac{n}{2}} & \rho^{\frac{n}{2}-1} & \cdots \rho 1 \rho \cdots & \rho^{\frac{n}{2}} \\ \cdot & & & & & \\ \cdot & & & & & \\ \rho^n & \rho^{n-1} & \cdots & \cdots & \rho & 1 \end{bmatrix} \begin{bmatrix} \phi_{00}^{(n)} \\ \phi_{01}^{(n)} \\ \cdot \\ \cdot \\ \cdot \\ \phi_{0\frac{n}{2}}^{(n)} \\ \cdot \\ \cdot \\ \phi_{00}^{(n)} \end{bmatrix}$$

$$= \begin{bmatrix} a_0^{(n)} & \phi_{00}^{(n)} \\ a_1^{(n)} & \phi_{01}^{(n)} \\ \cdot \\ \cdot \\ a_{\frac{n}{2}}^{(n)} & \phi_{0\frac{n}{2}}^{(n)} \\ \cdot \\ \cdot \\ a_0^{(n)} & \phi_{00}^{(n)} \end{bmatrix} \quad \text{where}$$

$$a_0(n) = \sigma^2 \left[1 + \rho \frac{\phi_{01}(n)}{\phi_{00}(n)} + \rho^2 \frac{\phi_{02}(n)}{\phi_{00}(n)} + \ldots + \rho^{n-1} \frac{\phi_{01}(n)}{\phi_{00}(n)} \right.$$

$$\left. + \rho^n \vphantom{\frac{}{}} \right]$$

$$\vdots$$

$$a_{\frac{n}{2}}(n) = \sigma^2 \left[1 + 2\rho \frac{\phi_{0\,\frac{n}{2}-1}(n)}{\phi_{0\,\frac{n}{2}}(n)} + 2\rho^2 \frac{\phi_{0\,\frac{n}{2}-2}(n)}{\phi_{0\,\frac{n}{2}}(n)} + \ldots \right.$$

$$\left. + 2\rho^{\frac{n}{2}-1} \frac{\phi_{01}(n)}{\phi_{0\,\frac{n}{2}}(n)} + 2\rho^{\frac{n}{2}} \frac{\phi_{00}(n)}{\phi_{0\,\frac{n}{2}}(n)} \right] .$$

But since $\phi_0(n)$ is an eigenvector $a_j(n) = \lambda_0(n)$,
$j = 0, \ldots, \frac{n}{2}$. Let $\tilde{\lambda}_0(n) = \frac{\lambda_0(n)}{\sigma^2}$, then since $\lambda_0(n) \to$

$f_x(0) = \frac{1}{(1-\rho)^2}$, $\tilde{\lambda}_0(n) \to \frac{1}{\sigma^2(1-\rho)^2} = \frac{(1-\rho^2)}{(1-\rho)^2}$

$$= \frac{(1-\rho)(1+\rho)}{(1-\rho)^2} = \frac{1+\rho}{1-\rho} = 1 + 2\rho + 2\rho^2 + \dots .$$ In the

expression for $a_{\frac{n}{2}}^{(n)}$ it follows from lemma 1 that the

$\frac{\frac{n}{2}}{\sigma^2}$

coefficient of each positive power of ρ is ≤ 2. Hence

in order for $a_{\frac{n}{2}}^{(n)}$ to approach $1 + 2\rho + 2\rho^2 + \dots$ it

$\frac{\frac{n}{2}}{\sigma^2}$

follows that for each fixed k, $\dfrac{\phi_0^{(n)} \frac{n}{2} - k}{\phi_0^{(n)} \frac{n}{2}} \rightarrow 1$. On the

other hand since $\sum_{\nu=0}^{n} \phi_{0\nu}^{(n)2} = 1$, then $\phi_0^{(n)} \frac{n}{2} \rightarrow 0$ as

$n \rightarrow \infty$.

<u>Theorem 2.1</u>: Let $\phi_0^{(n)}$ be as in the preceeding lemmas,

and let $g_0^{(n)}(\omega) = \sum_{\nu=0}^{n} \phi_{0\nu}^{(n)} e^{-i\nu\omega}$. Then $|g_0^{(n)}(\omega)|^2$

$\rightarrow \delta(\omega)$ as $n \rightarrow \infty$.

<u>Proof</u>: $|g_0^{(n)}(\omega)|^2 = \sum_{k=-n}^{n} c_k e^{ik\omega}$, where $c_k = c_{-k}$

$= \sum_{\nu=k}^{n} \phi_{0\nu}^{(n)} \phi_{0\nu-k}^{(n)}$ so that in particular $c_0 = \sum_{\nu=0}^{n} \phi_{0\nu}^{(n)2}$

=1. For $k < \frac{n}{2}$, $\sum_{j=0}^{n} \phi_{0j}^{(n)^2} - \sum_{j=k}^{n} \phi_{0j}^{(n)} \phi_{0j-k}^{(n)}$

$= \phi_{00}^{(n)^2} + \ldots + \phi_{0k-1}^{(n)^2} + \phi_{0k}^{(n)}(\phi_{0k}^{(n)} - \phi_{00}^{(n)})$

$+ \phi_{0k+1}^{(n)}(\phi_{0k+1} - \phi_{01}) + \ldots + \phi_{0\frac{n}{2}}(\phi_{0\frac{n}{2}} - \phi_{0\frac{n}{2}-k})$

$+ \phi_{0\frac{n}{2}+1}(\phi_{0\frac{n}{2}+1} - \phi_{0\frac{n}{2}+1-k}) + \ldots + \phi_{0 n-1}(\phi_{0 n-1} - \phi_{0 n-k-1})$

$+ \phi_{0n}(\phi_{0n} - \phi_{0n-k})$. But $\phi_{00} = \phi_{0n}$, $\phi_{01} = \phi_{0 n-1}$,

$\ldots$, $\phi_{0\frac{n}{2}-1} = \phi_{0\frac{n}{2}+1}$, implies $\phi_{0k}(\phi_{0k} - \phi_{00})$

$+ \phi_{0n}(\phi_{0n} - \phi_{0n-k}) = \phi_{0k}(\phi_{0k} - \phi_{00}) - \phi_{00}(\phi_{0k} - \phi_{00})$

$= (\phi_{0k} - \phi_{00})^2$, etc., so that the above sum becomes

$$\sum_{j=0}^{n} \phi_{0j}^2 - \sum_{j=k}^{n} \phi_{0j} \phi_{0j-k} = \sum_{j=0}^{k-1} \phi_{0j}^2 + \sum_{j=k}^{n/2} (\phi_{0j} - \phi_{0j-k})^2$$

$= S_1 + S_2 > 0$. Since $\phi_{0\frac{n}{2}}^{(n)} \to 0$, then for fixed k,

$S_1 \leq k(\phi_{0\frac{n}{2}}^{(n)})^2 \to 0$ as $n \to \infty$. Let r be the greatest in-

teger such that $rk \leq \frac{n}{2}$, and let $a_1 = \min(rk+1, \frac{n}{2})$, a_2
$= \min(rk+2, \frac{n}{2}), \ldots, a_{k-1} = \min((r+1)k-1, \frac{n}{2})$. Then S_2

$$= \sum_{j=k}^{n/2} (\phi_{0j} - \phi_{0,j-k})^2 \leq (\phi_{0k} - \phi_{00})^2 + (\phi_{02k} - \phi_{0k})^2 + \ldots$$

$$+ (\phi_{0\ rk} - \phi_{0(r-1)k})^2 + (\phi_{0\ \frac{n}{2}} - \phi_{0\ rk})^2 + (\phi_{0\ k+1} - \phi_{01})^2$$

$$+ \ldots + (\phi_{0\ a_1} - \phi_{0\ (r-1)k+1})^2 + (\phi_{0\ \frac{n}{2}} - \phi_{0\ a_1})^2 \ldots$$

$$+ (\phi_{0\ 2k-1} - \phi_{0\ k-1})^2 + \ldots + (\phi_{0\ a_{k-1}} - \phi_{0\ rk-1})^2$$

$$+ (\phi_{0\ \frac{n}{2}} - \phi_{0\ a_{k-1}})^2.$$ Now, each term of $(\phi_{0k} - \phi_{00})$

$$+ (\phi_{0\ 2k} - \phi_{0k}) + \ldots + (\phi_{0\ rk} - \phi_{0\ (r-1)k}) + (\phi_{0\ \frac{n}{2}} - \phi_{0\ rk})$$

$$= (\phi_{0\ \frac{n}{2}} - \phi_{00})$$ is non-negative so that the sum of squares

of these terms is less than or equal to the square of the

sum. The same is true for each of the other $k-1$ sums so

that $S_2 \leq k(\phi_{0\ \frac{n}{2}} - \phi_{00})^2 \to 0$ as $n \to \infty$. But this implies

that for each fixed k, $c_k = c_{-k} \to 1$ as $n \to \infty$. From this

it follows that for any positive integer N, and any $\epsilon > 0$, there exists an integer M such that $n > M$

$$\Rightarrow |1 - c_k^{(n)}| < \frac{\epsilon}{2N+1} \text{ for } k = 0, 1, \ldots, N. \text{ Then}$$

$$|\sum_{k=-N}^{N} c_k^{(n)} e^{ik\omega} - \sum_{k=-N}^{N} e^{ik\omega}| \leq \sum_{k=-N}^{N} |c_k^{(n)} - 1| < \epsilon.$$

Thus, if the partial sums $S_N(\omega) = \sum_{k=-N}^{N} e^{ik\omega}$ of the

trigonometrical series $\sum_{k=-\infty}^{\infty} e^{ik\omega}$ converge to a general-

ized function $g(\omega)$, the transfer function $|g_0^{(n)}(\omega)|^2$

$= \sum_{k=-n}^{n} c_k^{(n)} e^{ik\omega}$ must converge to the same function.

It is shown, however, in Lighthill [6, p. 67] that the

trigonometrical series $\sum_{k=-\infty}^{\infty} e^{ik\omega}$ converges to the gen-

eralized function $g(\omega) = 2\Pi \sum_{k=-\infty}^{\infty} (\omega - 2k\Pi)$, (a train of

delta functions). When ω is restricted to the interval $[-\Pi, \Pi]$, $g(\omega)$ is simply $2\Pi\delta(\omega)$, and it then follows that

$$|g_0^{(n)}(\omega)|^2 \xrightarrow{n \to \infty} 2\Pi\delta(\omega).$$

Now suppose $\rho < 0$. Let the covariance matrices and spectral density functions now be denoted by $R^{(n)-}$ and $f_x^-(\omega)$. Since $f_x^-(\omega) = f_x(\omega + \Pi)$, it follows from the theory of Toeplitz forms that the eigenvalues of $R^{(n)-}$ are the same as those of $R^{(n)}$.

Theorem 2.2: Let $g_0^{(n)-}(\omega) = \sum_{\nu=0}^{n} \phi_{0\nu}^{(n)-} e^{-i\nu\omega}$, where $\phi_{0\nu}^{(n)-}$ is the νth element of $\phi_0^{(n)-}$, the eigenvector corresponding to the largest eigenvalue $\lambda_0^{(n)}$ of $R^{(n)-}$. Then $g_0^{(n)-}(\omega) \to 0$ as $n \to \infty$ for any fixed ω such that $|\omega| \neq \Pi$.

Proof: By direct multiplication, it may be verified that $\phi_{0\nu}^{(n)-} = (-1)^{\nu} \phi_{0\nu}^{(n)}$ so that

$$g_0^{(n)-}(\omega) = \sum_{\nu=0}^{n} (-1)^{\nu} \phi_{0\nu}^{(n)} e^{-i\nu\omega} = \sum_{\nu=0}^{n} \phi_{0\nu}^{(n)} e^{-i(\omega\pm\Pi)\nu} \to 0$$

whenever $(\omega \pm \Pi) \neq 0$.

These same results can be obtained to show the behavior of the first principal component process of certain higher order moving average and autoregressive processes. For instance, consider the second order moving average process $x(t) = Z(t) + \alpha_1 Z(t - 1) + \alpha_2 Z(t - 2)$, where $Z(t)$ is a white noise process, $0 < \alpha_1 < 1$, $0 < \alpha_2 < 1$. $x(t)$ has spectral density function $f_x(\omega) = \frac{1}{2\pi} (1 + \alpha_1^2 + \alpha_2^2 + 2(\alpha_1 + \alpha_1\alpha_2)\cos\omega + 2\alpha_2\cos2\omega)$ which attains its maximum at $\omega = 0$. The covariance matrix in this case is

$$R^{(n)} = \begin{bmatrix} 1+\alpha_1^2+\alpha_2^2 & \alpha_1+\alpha_1\alpha_2 & \alpha_2 & 0 & \cdots & \cdots & 0 \\ \alpha_1+\alpha_1\alpha_2 & 1+\alpha_1^2+\alpha_2^2 & \alpha_1+\alpha_1\alpha_2 & \alpha_2 & 0 & \cdots & 0 \\ \alpha_2 & \alpha_1+\alpha_1\alpha_2 & 1+\alpha_1^2+\alpha_2^2 & \alpha_1+\alpha_1\alpha_2 & \alpha_2 & & 0\cdots 0 \\ 0 & & & & & & \\ \vdots & & & & & & \\ 0 & \cdots & \cdots & & \alpha_2 & \alpha_1+\alpha_1\alpha_2 & 1+\alpha_1^2+\alpha_2^2 \end{bmatrix}$$

The results of Perron's theorem and its corollary can

be shown (Frobenius' theorem [h]) to hold for an irreducible matrix $A = (a_{ij})$, where A is called irreducible if $a_{ij} \geq 0$ and if whenever $a_{pq} = 0$ for $p \neq q$, there exists some set of indices $i_1, \ldots, i_r$ such that the indices $p, i_1, \ldots, i_r, q$ are all different and $a_{pi_1} \neq 0, \ldots, a_{i_{r-1}k_r} \neq 0$ $a_{i_r q} \neq 0$. It can easily be verified that $R^{(n)}$ is irreducible, for example, $a_{1\ n+1} = 0$ but $a_{13} \neq 0$, $a_{35} \neq 0$, $\ldots a_{n-1\ n+1} \neq 0$.

Lemma 2.1 holds by exactly the same proof as in the autoregressive case. To show that lemma 2.2 holds, consider $R^{(n)} \phi_0^{(n)}$

$$= \begin{bmatrix} \alpha & \beta & \gamma & 0 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 0 \\ \beta & \alpha & \beta & \gamma & 0 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 0 \\ \gamma & \beta & \alpha & \beta & \gamma & 0 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 0 \\ 0 & \gamma & \beta & \alpha & \beta & \gamma & 0 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 0 \\ \cdot & & & & & & & & & & & & & & & & & \\ \cdot & & & & & & & & & & & & & & & & & \\ 0 & \cdot & \cdot & \cdot & \cdot & \cdot & \gamma & \beta & \alpha & \beta & \gamma & 0 & \cdot & \cdot & \cdot & \cdot & & 0 \\ 0 & \cdot & \cdot & \cdot & \cdot & \cdot & 0 & \gamma & \beta & \alpha & \beta & \gamma & 0 & \cdot & \cdot & \cdot & \cdot & 0 \\ \cdot & & & & & & & & & & & & & & & & & \\ \cdot & & & & & & & & & & & & & & & & & \\ \cdot & & & & & & & & & & & & & & & & & \end{bmatrix} \begin{bmatrix} \phi_0 \\ \phi_1 \\ \phi_2 \\ \cdot \\ \cdot \\ \phi_{\frac{n}{2}-1} \\ \phi_{\frac{n}{2}} \\ \cdot \\ \cdot \\ \phi_0 \end{bmatrix}$$

$$= \begin{bmatrix} a_0^{(n)} & \phi_0 \\ a_1^{(n)} & \phi_1 \\ \cdot \\ \cdot \\ \cdot \\ a_{\frac{n}{2}-2}^{(n)} & \phi_{\frac{n}{2}-2} \\ a_{\frac{n}{2}-1}^{(n)} & \phi_{\frac{n}{2}-1} \\ a_{\frac{n}{2}}^{(n)} & \phi_{\frac{n}{2}} \\ \cdot \\ \cdot \\ \cdot \end{bmatrix} \quad ,$$

where $a_0^{(n)} = \alpha + \beta \dfrac{\phi_1}{\phi_0} + \gamma \dfrac{\phi_2}{\phi_0}$

$a_1^{(n)} = \alpha + \beta \left(\dfrac{\phi_0}{\phi_1} + \dfrac{\phi_2}{\phi_1} \right) + \gamma \left(\dfrac{\phi_2}{\phi_1} \right)$

$\cdot$
$\cdot$
$\cdot$

$a_{\frac{n}{2}-2}^{(n)} = \alpha + \beta \left[\dfrac{\phi_{\frac{n}{2}-1}}{\phi_{\frac{n}{2}-2}} + \dfrac{\phi_{\frac{n}{2}-3}}{\phi_{\frac{n}{2}-2}} \right] + \gamma \left[\dfrac{\phi_{\frac{n}{2}}}{\phi_{\frac{n}{2}-2}} + \dfrac{\phi_{\frac{n}{2}-4}}{\phi_{\frac{n}{2}-2}} \right]$

$a_{\frac{n}{2}-1}^{(n)} = \alpha + \beta \left[\dfrac{\phi_{\frac{n}{2}}}{\phi_{\frac{n}{2}-1}} + \dfrac{\phi_{\frac{n}{2}-2}}{\phi_{\frac{n}{2}-1}} \right] + \gamma \left[1 + \dfrac{\phi_{\frac{n}{2}-3}}{\phi_{\frac{n}{2}-1}} \right]$

$a_{\frac{n}{2}}^{(n)} = \alpha + 2\beta \dfrac{\phi_{\frac{n}{2}-1}}{\phi_{\frac{n}{2}}} + 2\gamma \dfrac{\phi_{\frac{n}{2}-2}}{\phi_{\frac{n}{2}}} \quad ,$

and $\alpha = 1 + \alpha_1{}^2 + \alpha_2{}^2$, $\beta = \alpha_1 + \alpha_1\alpha_2$, $\gamma = \alpha_1$. Since ϕ_0 is an eigenvector, $a_0{}^{(n)} = a_1{}^{(n)} = \ldots = a_{\frac{n}{2}}{}^{(n)}$

$= \lambda_0{}^{(n)}$ and as $n \to \infty$, $\lambda_0{}^{(n)} \to f_x(0) = \alpha + 2\beta + 2\gamma$.

The expression for $a_{\frac{n}{2}}{}^{(n)}$ thus shows that as $n \to \infty$,

$\phi_{\frac{n}{2}-1}/\phi_{\frac{n}{2}} \to 1$ and $\phi_{\frac{n}{2}-2}/\phi_{\frac{n}{2}} \to 1$. Working backwards then,

it follows that for any fixed k, $\phi_{\frac{n}{2}-k}/\phi_{\frac{n}{2}} \to 1$ as $n \to \infty$

and the lemma follows. Theorem 2.1 then holds exactly as before. If $\alpha_1 < 0$, $\alpha_2 > 0$, then $\alpha_1 + \alpha_1\alpha_2 < 0$ and the spectral density function now denoted by $f_x^-(\omega)$ assumes its maximum at $\lambda = \pm\Pi$. $f_x^-(\omega) = f_x(\omega + \Pi)$ so

that the eigenvalues of $R^{(n)-}$ are the same as those of $R^{(n)}$ except with different ordering, where $R^{(n)-}$ is the covariance matrix in the case with $\alpha_1 < 0$. $R^{(n)-}$

can be derived from $R^{(n)}$ by changing the algebraic sign of every other element and, hence, direct multiplication

will show that $\phi_0{}^{(n)-} = \begin{pmatrix} \phi_{00}{}^{(n)-} \\ \cdot \\ \cdot \\ \phi_{0n}{}^{(n)-} \end{pmatrix}$, where $\phi_{0\nu}{}^{(n)-}$

$= (-1)^{\nu} \phi_{0\nu}^{(n)}$, so that theorem 2.2 holds.

Mixed Autoregressive Moving Average Process

Let $x(t)$ be a stationary process of the form

$$x(t) - \rho x(t-1) = Z(t) + \alpha Z(t-1). \qquad (2.13)$$

This can be formally written as,

$$(1 - \rho B)x(t) = (1 + \alpha B)Z(t)$$

or,

$$x(t) = \frac{(1 + \alpha B)}{(1 - \rho B)} \, Z(t), \text{ where B denotes} \qquad (2.14)$$

the backwards lag operator. From (2.14) if $|\rho| < 1$, it follows that $x(t)$ can be expressed as the infinite moving average process

$$x(t) = Z(t) + (\rho + \alpha)Z(t-1) + \rho(\rho + \alpha)Z(t-2) + \ldots$$

from which it follows that

$$r(0) = \frac{1 + 2\rho\alpha + \alpha^2}{1 - \rho^2} = \frac{\delta}{1-\rho^2} \, ,$$

$$r(n) = \frac{\rho^{n-1}[\alpha + \rho(\alpha^2 + \rho\alpha + 1)]}{1 - \rho^2} = \frac{\rho^{n-1}\gamma}{1 - \rho^2} \, ,$$

so that,

$$R^{(n)} = \begin{bmatrix} \delta & \gamma & \rho\gamma & \rho^2\gamma & \cdots & \rho^{n-1}\gamma \\ \gamma & \delta & \gamma & \rho\gamma & \cdots & \rho^{n-2}\gamma \\ \vdots & & & & & \\ \vdots & & & & & \\ \rho^{n-1}\gamma & & \cdots & & \gamma & \delta \end{bmatrix} \frac{1}{1-\rho^2}$$

If $\rho, \alpha > 0$, $R^{(n)}$ is a positive matrix and if

$S_0^{(1)}, S_1^{(1)}, \ldots, S_n^{(1)}$ denote its row sums, then

$S_0^{(1)} < S_1^{(1)} < \ldots < S_{\frac{n}{2}}^{(1)} > \ldots > S_n^{(1)} = S_0^{(1)}$. $S_0^{(2)}$

$= \delta S_0^{(1)} + \gamma S_1^{(1)} + \rho\gamma S_2^{(1)} + \ldots + \rho^{\frac{n}{2}-1}\gamma S_{\frac{n}{2}}^{(1)} + \ldots$

$+ \rho^{n-1}\delta S_n^{(1)}$, $S_1^{(2)} = \gamma S_0^{(1)} + \delta S_1^{(1)} + \gamma S_2^{(1)} + \ldots$

$+ \rho^{\frac{n}{2}-1}\gamma S_{\frac{n}{2}}^{(1)} + \ldots + \rho^{n-2}\delta S_n^{(1)}$. $\gamma < \delta$ since it is al-

ways true that $|r(t)| < r(0)$ and γ is real. Hence, it

follows that $S_1^{(2)} > S_0^{(2)}$, and in a similar way $S_{\frac{n}{2}}^{(2)}$

$> S_{\frac{n}{2}-1}^{(2)} > \ldots > S_1^{(2)} > S_0^{(2)}$, etc., which implies that

lemma 2.1 holds. The spectral density function of $x(t)$

$$f_x(\omega) = \frac{1}{2\pi} \left| \frac{1 + \alpha e^{-i\omega}}{1 - \rho e^{-i\omega}} \right|^2 = \frac{1}{2\pi} \frac{1 + \alpha^2 + 2\alpha\cos\omega}{1 + \rho^2 - 2\rho\cos\omega}.$$

For fixed n, the row sum $S_{\frac{n}{2}}^{(1)}$

$$= (\delta + 2\gamma + 2\rho\gamma + \ldots + 2\rho^{\frac{n}{2}-1}\gamma) \frac{1}{1 - \rho^2}$$

$$= \frac{1}{1 - \rho^2} [1 + 2\rho\alpha + \alpha^2 + 2(\alpha + \rho(\alpha^2 + \rho\alpha + 1))$$

$$+ \rho(\alpha + \rho(\alpha^2 + \rho\alpha + 1)) + \ldots + 2\rho^{\frac{n}{2}-1}(\alpha + \rho(\alpha^2 + \rho\alpha + 1))]$$

$$= \frac{1}{1 - \rho^2} [(1 + \alpha)^2 + 2\rho(1 + \alpha)^2 + \ldots + 2\rho^{\frac{n}{2}-1}(1 + \alpha)^2$$

$$+ 2\rho^{\frac{n}{2}}(\alpha^2 + \alpha + 1) + 2\rho^{\frac{n}{2}+1}\alpha]. \text{ As } n \to \infty \text{ this becomes}$$

$$\frac{1}{1 - \rho^2} (1 + \alpha)^2(\frac{2}{1 - \rho} - 1) \quad = \frac{1}{1 - \rho^2} (1+\alpha)^2 \frac{(1 + \rho)}{(1 - \rho)}$$

$$= \frac{(1 + \alpha)^2}{(1 - \rho)^2} = 2\Pi f_x(0).$$ The same argument as in the auto-

regressive case then shows that lemma 2.2 and hence

theorem 2.1 also hold in this case.

If both $\rho < 0$ and $\alpha < 0$ then $\delta > 0$, while $\gamma < 0$;
$f_x^-(\omega) = f_x(\Pi + \omega)$ has its maximum at $\omega = \pm\Pi$; $\phi_\theta^{(n)-}$

$$= \begin{bmatrix} \phi_{00}^{(n)-} \\ \cdot \\ \cdot \\ \cdot \\ \phi_{0n}^{(n)-} \end{bmatrix}, \text{ where } \phi_{0\nu}^{(n)-} = (-1)^\nu \phi_{0\nu}^{(n)} \text{ and hence}$$

theorem 2.2 holds.

CHAPTER III

PROCFSSES WITH NONCONSTANT MEANS

Let $y(t)$ be a real valued second order sta-
tionary process with mean zero, and let $x(t) = m(t)$
$+ y(t)$, where $m(t)$ is a deterministic but unknown func-
tion. In this chapter the covariances considered will
be expected sample covariances where the sample covari-
ances are computed as if $m(t)$ were constant. Thus, each
covariance will consist of a part due to $m(t)$ and a part
due to $y(t)$ as follows.

Let $x(1)$, $x(2),\ldots,$ $x(n + k)$ be a sample from
the process $x(t)$ and define

$$\bar{x}^{(r)} = \frac{1}{n} \sum_{t=k+1}^{n+k} x(t - r) = \bar{m}^{(r)} + \bar{y}^{(r)}, \quad r = 0,\ldots,k$$

$$(3.1)$$

The sample covariance between $x(t)$ and $x(t - \tau)$ can
then be written as

$$S(\tau) = \frac{1}{n} \sum_{t=k+1}^{n+k} (x(t) - \bar{x}^{(0)})(x(t - \tau) - \bar{x}^{(\tau)})$$

55

$$= \frac{1}{n} \sum_{t=k+1}^{n+k} (m(t) + y(t) - \overline{m}^{(0)} - \overline{y}^{(0)})(m(t - \tau)$$

$$+ y(t - \tau) - \overline{m}^{(\tau)} - \overline{y}^{(\tau)})$$

$$\bullet \quad \frac{1}{n} \sum_{t=k+1}^{n+k} m(t)m(t - \tau) - \overline{m}^{(0)}\overline{m}^{(\tau)},$$

$$+ \frac{1}{n} \sum_{t=k+1}^{n+k} m(t)[y(t - \tau) - \overline{y}^{(\tau)}]$$

$$+ \frac{1}{n} \sum_{t=k+1}^{n+k} m(t - \tau)[y(t) - \overline{y}^{(0)}]$$

$$+ \frac{1}{n} \sum_{t=k+1}^{n+k} y(t)y(t - \tau) - \overline{y}^{(0)}\overline{y}^{(\tau)} \quad .$$

Since $y(t)$ is stationary then (approximately) $ES(\tau)$

$$= \frac{1}{n} \sum_{t=k+1}^{n+k} m(t)m(t - \tau) - \overline{m}^{(0)}\overline{m}^{(\tau)} \pm r_y(\tau), \quad (3.2)$$

If $y(t)$ is a white noise process, then the equation becomes

$$E\ S(\tau) = \frac{1}{n} \sum_{t=k+1}^{n+k} m(t)m(t - \tau) - \overline{m}^{(0)}\overline{m}^{(\tau)} + \sigma_y^2 \delta_{\tau 0}$$

$$(3.3)$$

x(t) A Linear Trend Plus White Noise

Let $x(t) = m(t) + Z(t)$, where $Z(t)$ is a
white noise process, $EZ(t) = 0$, $E(Z(t) \, Z(t - \tau))$
$= \sigma_Z^2 \delta_{0\tau}$, and $m(t) = a + bt$ where a and b are constants.
Then $m(t) \, m(t - \tau) = (a + bt)(a + b(t - \tau))$
$= (a + bt)((a - b\tau) + bt) = aa_\tau + (ab + a_\tau b)t + b^2 t^2$

$$\frac{1}{n} \sum_{t=k+1}^{n+k} m(t) \, m(t - \tau) = aa_\tau + \frac{b(a + a_\tau)}{n} \sum_{t=k+1}^{n+k} t$$

$$+ \frac{b^2}{n} \sum_{t=k+1}^{n+k} t^2, \text{ while } \overline{m}(0)\overline{m}(\tau)$$

$$= \frac{1}{n} \sum_{t=k+1}^{n+k} (a + bt) \frac{1}{n} \sum_{t=k+1}^{n+k} (a + b(t - \tau))$$

$$= (a + \frac{b}{n} \sum_{t=k+1}^{n+k} t)(a_\tau + \frac{b}{n} \sum_{t=k+1}^{n+k} t)$$

.

$$= aa_\tau + \frac{b(a + a_\tau)}{n} \sum_{t=k+1}^{n+k} t + \frac{b^2}{n^2} [\sum_{t=k+1}^{n+k} t]^2$$

By (3.3) $ES_\tau = \frac{b^2}{n} \sum_{t=k+1}^{n+k} t^2 - \frac{1}{n}[\sum_{t=k+1}^{n+k} t]^2 + \sigma_Z^2 \delta_{\tau 0}.$

Hence, the expected covariance matrix $ER^{(n)}$ has the form

$$ER^{(n)} = \begin{bmatrix} \sigma^2+\alpha & \alpha_2 \cdot & \cdots & \cdots & \alpha \\ \alpha & \sigma^2+\alpha & \alpha & \cdots & \alpha \\ \cdot & & & & \\ & & & & \\ \alpha & \cdots & \cdots & \alpha & \sigma^2+\alpha \end{bmatrix} , \text{ where}$$

α = the first term above. This is a positive circular, matrix, and hence for a given n, $ER^{(n)}$ has eigenvalues

$$\lambda_\nu^{(n)} = \sigma^2 + \alpha + (\sum_{j=1}^{n} r_\nu^j) \alpha, \text{ where } r_\nu \text{ is a root of}$$

$x^{n+1} - 1 = 0$. When $r_\nu = 1$, the corresponding eigenvalue is $\lambda_0^{(n)} = \sigma^2 + (n + 1)\alpha$, while if $r_\nu \neq 1$, $\lambda_\nu^{(n)} = \sigma^2$.

The eigenvector corresponding to $\lambda_0^{(n)}$ is $\phi_0^{(n)} = \begin{bmatrix} \frac{1}{\sqrt{n+1}} \\ \cdot \\ \cdot \\ \cdot \\ \frac{1}{\sqrt{n+1}} \end{bmatrix}$,

whose transfer function is $|g_0^{(n)}(\omega)|^2$

$$= \left| \frac{1}{\sqrt{n+1}} \sum_{k=0}^{n} e^{-ik\omega} \right|^2 . \quad \text{But,} \frac{1}{\sqrt{n+1}} \sum_{k=0}^{n} (e^{-i\omega})^k$$

$$= \frac{1}{\sqrt{n+1}} \left[\frac{1 - e^{-i\omega(n+1)}}{1 - e^{i\omega}} \right] \xrightarrow{n\to\infty} 0 \text{ for any fixed } \omega \neq 0.$$

On the other hand since $m(t)$ is non-periodic (a linear trend in this case) it can be considered as having the Fourier representation

$$m(t) = \int_{-\Pi}^{\Pi} e^{it\omega} \, d\mu(\omega),$$

where $d\mu(\omega) = 0$ for $\omega \neq 0$. As a result, the (sample) power spectrum of $x(t)$ consists of the spectrum of $y(t)$ plus a jump at $\omega = 0$ due to $m(t)$. The form of the above transfer function shows that the first principal component tends to remove this zero frequency or trend portion of $x(t)$

$x(t)$ A Periodic Function Plus White Noise

Consider the process $x(t) = \cos(\frac{\Pi}{2}t) + Z(t)$, where $Z(t)$ is a white noise process. For convenience, it will be assumed that both n (the sample size) and $N + 1$ (the order of the covariance matrix) are always taken to be divisible by four. Then,

$$\overline{m}^{(0)} = \frac{1}{n} \sum_{t=k+1}^{n+k} \cos(\frac{\Pi}{2}t) = 0,$$

$$\frac{1}{n} \ \Sigma m(t)m(t) = \frac{1}{n} (0 + 1 + \ldots) = 1/2,$$

$$\frac{1}{n} \ \Sigma m(t)m(t - 1) = 0,$$

$$\frac{1}{n} \ \Sigma m(t)m(t - 2) = \frac{1}{n} (0 - 1 + 0 \ldots) = -1/2$$

$$\frac{1}{n} \ \Sigma m(t)m(t - 3) = 0, \text{ etc.}$$

Hence, the expected covariance matrix is of the form

$$E \ R^{(n)} = \begin{bmatrix} \sigma_z^2 + 1/2 & 0 & -1/2 & 0 & \ldots & 1/2 & 0 & -1/2 & 0 \\ 0 & & \sigma_z^2 + 1/2 & 0 & \ldots\ldots\ldots\ldots\ldots\ldots & -1/2 \\ \cdot \\ \cdot \\ \cdot \\ 0 & \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots & 0 \cdot \sigma_z^2 + 1/2 \end{bmatrix}$$

This, again, is a circular matrix with eigenvalues

$$\lambda_\nu = (\sigma_z^2 + 1/2) + 1/2 \sum_{j=1}^{\frac{N+1}{2}} (-1)^j r_\nu^{2j}, \text{ where } r_\nu \text{ is}$$

a root of $x^{N+1} - 1 = 0$. The expression for λ_ν assumes

its maximum value when $(r_\nu)^{2j} = (-1)^j$ or when $r_\nu = \pm 1$.

Since r_ν can be expressed in the form $r_\nu = e^{i\frac{2\pi\nu}{N+1}}$,

$\nu = 0,1,\ldots,N$. $\lambda(\text{max.})$ is attained for $\nu = \frac{N+1}{4}$ and

$\nu = \frac{3(N+1)}{4}$. Using the same notation as in Chapter II,

two orthogonal eigenvectors corresponding to the double

root $\lambda(\text{max.})$ are given by

$$\phi_0^{(N)} = \frac{1}{\sqrt{N+1}} \begin{bmatrix} 1 \\ e^{i\frac{\Pi}{2}} \\ \cdot \\ \cdot \\ \cdot \\ e^{i\frac{\Pi N}{2}} \end{bmatrix} \quad \text{and} \quad \phi_0^{(N)} = \frac{1}{\sqrt{N+1}} \begin{bmatrix} 1 \\ e^{i\frac{3\Pi}{2}} \\ \cdot \\ \cdot \\ \cdot \\ e^{i\frac{3\Pi}{2}N} \end{bmatrix} \quad .$$

Let $\psi^{(N)}$ be an arbitrary linear combination $\psi^{(N)}$ $= a\phi_0^{(N)} + b\phi_0^{(N)}$. The transfer function corresponding to $\psi^{(N)}$ is

$$|g^{(N)}(\omega)|^2 = |\sum_{k=0}^{N} (\frac{a}{\sqrt{N+1}} e^{i\frac{\Pi}{2}k} + \frac{b}{\sqrt{N+1}} e^{i3\frac{\Pi}{2}k}) e^{-ik\omega}|^2 ,$$

where $g^{(N)}(\omega) = \frac{a}{\sqrt{N+1}} \sum_{k=0}^{N} e^{i(\frac{\Pi}{2}-\omega)k} + \frac{b}{\sqrt{N+1}} \sum e^{i(\frac{3\Pi}{2}-\omega)k} \rightarrow 0$

with N for any value of ω different from $\Pi/2$ or $3\Pi/2$. For $t = 0,1,\ldots,$ $\cos(\frac{\Pi}{2})t = \cos(\frac{3\Pi}{2})t$, so that any principal component formed from an eigenvector corresponding to $\lambda(max.)$ will tend to filter out the deterministic function $m(t)$.

APPENDIX

Figures 1 and 2 of chapter II showed the transfer functions corresponding to the two eigenvectors of the 2 X 2, (n = 1) covariance matrix of the moving average process $x(t) = Z(t) + .9Z(t - 1)$. Figures 3 and 4 show how these transfer functions change as n increases. Figure 3 shows the transfer functions corresponding to the eigenvectors ϕ_0, ϕ_2, ϕ_4 when n = 4, while figure 4 shows the transfer functions corresponding to ϕ_0, ϕ_3, ϕ_6 for n = 6.

$|g_0^{(4)}(\omega)|^2$

$|g_2^{(4)}(\omega)|^2$

$|g_4^{(4)}(\omega)|^2$

6.0

5.0

4.0

3.0

2.0

1.0

$\frac{\pi}{8}$ $\frac{\pi}{4}$ $\frac{3\pi}{8}$ $\frac{\pi}{2}$ $\frac{5\pi}{8}$ $\frac{3\pi}{4}$ $\frac{7\pi}{8}$ π

Figure 3

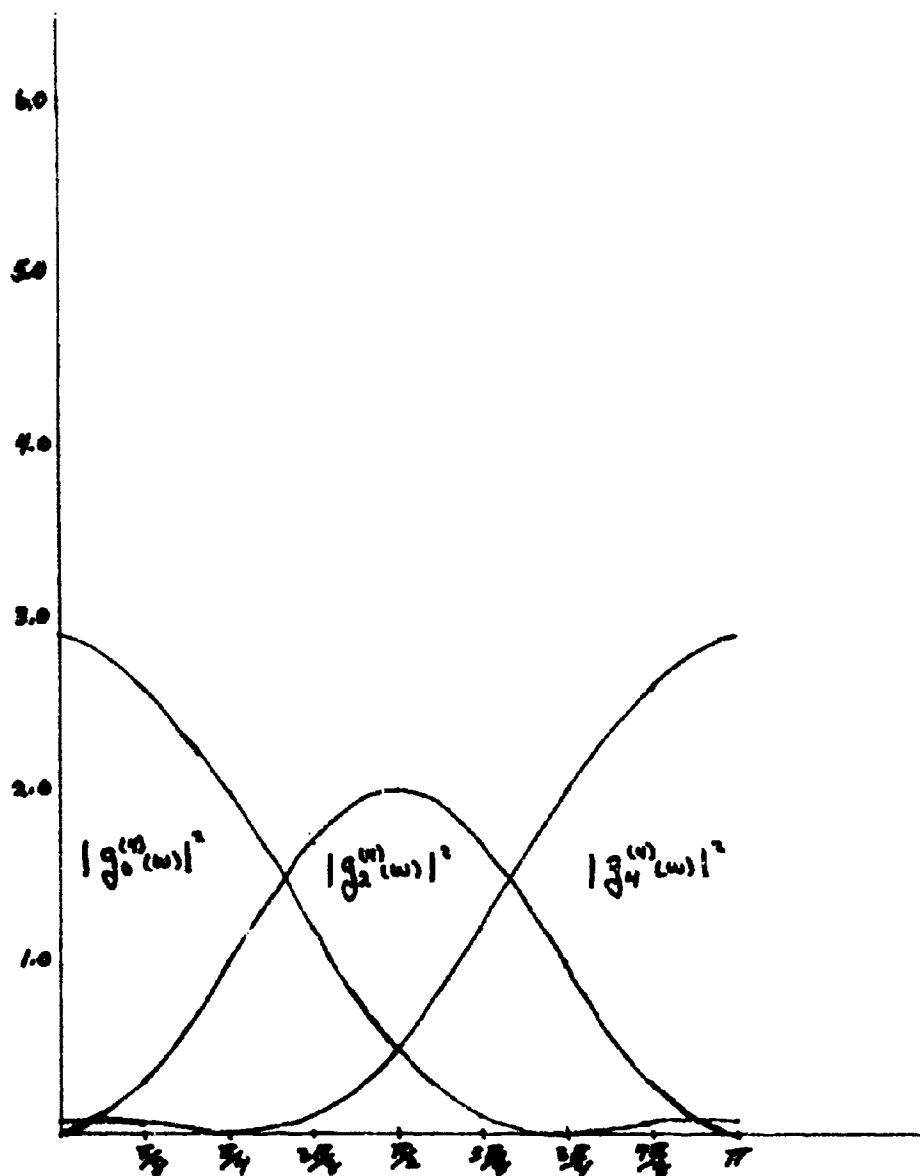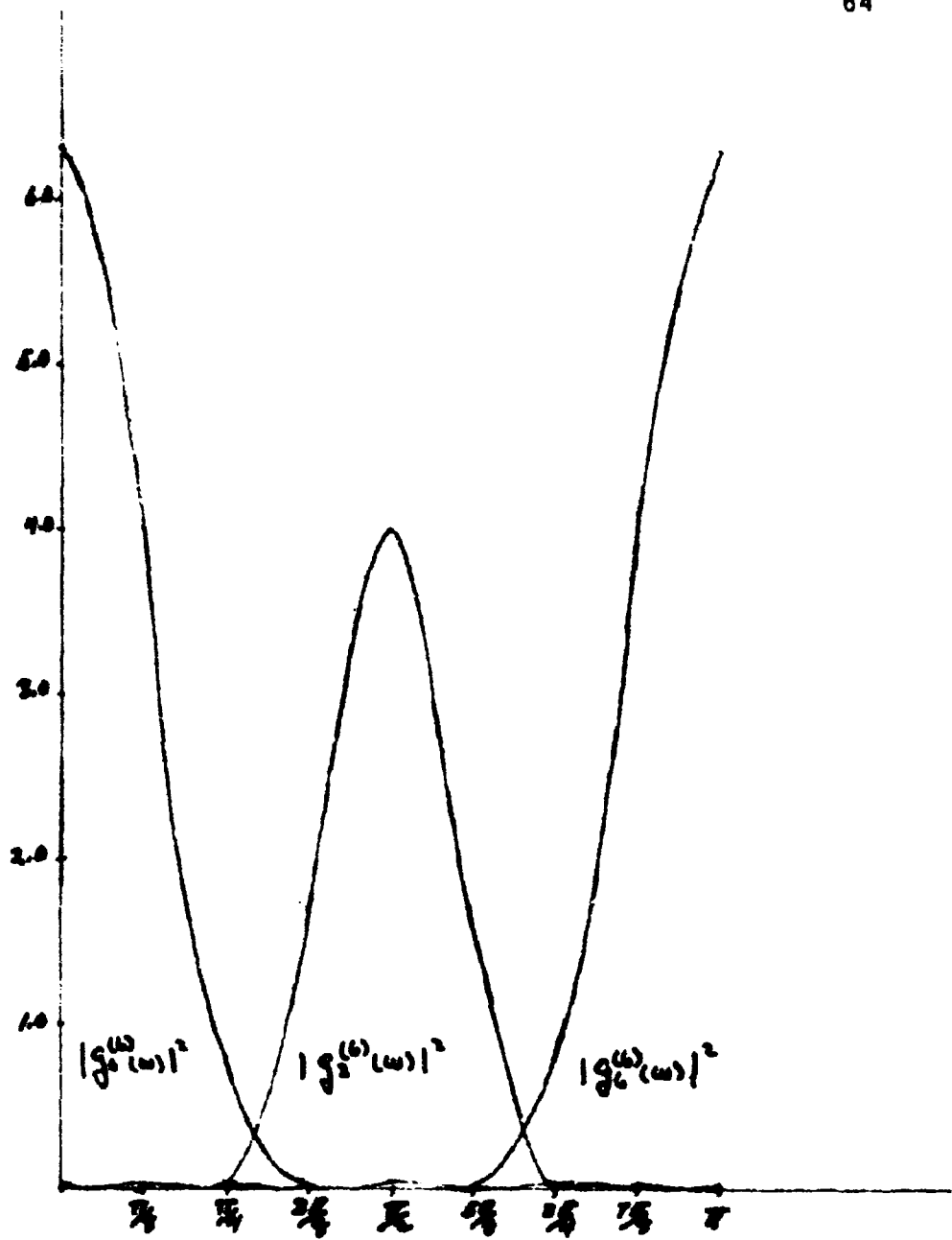$|g_0^{(6)}(\omega)|^2$ $|g_2^{(6)}(\omega)|^2$ $|g_6^{(6)}(\omega)|^2$

Figure 4

Figure 5 shows an example of the behavior
of an autoregressive process $x(t) + .9x(t - 1) = Z(t)$
together with its first and fifth principal component
processes for the case $n = 4$. The series $Z(t)$ consists
of a series of random numbers taken from the Rand table
of Normal Random Deviates. From this series, the series
$y(t)$ was constructed. The principal component processes
were constructed by numerically determining the eigen-
vectors of the theoretical covariance matrix of $y(t)$
and applying the appropriate filters to $y(t)$. The first
principal component process can be seen to consist pri-
marily of oscillations with period two, while the fifth
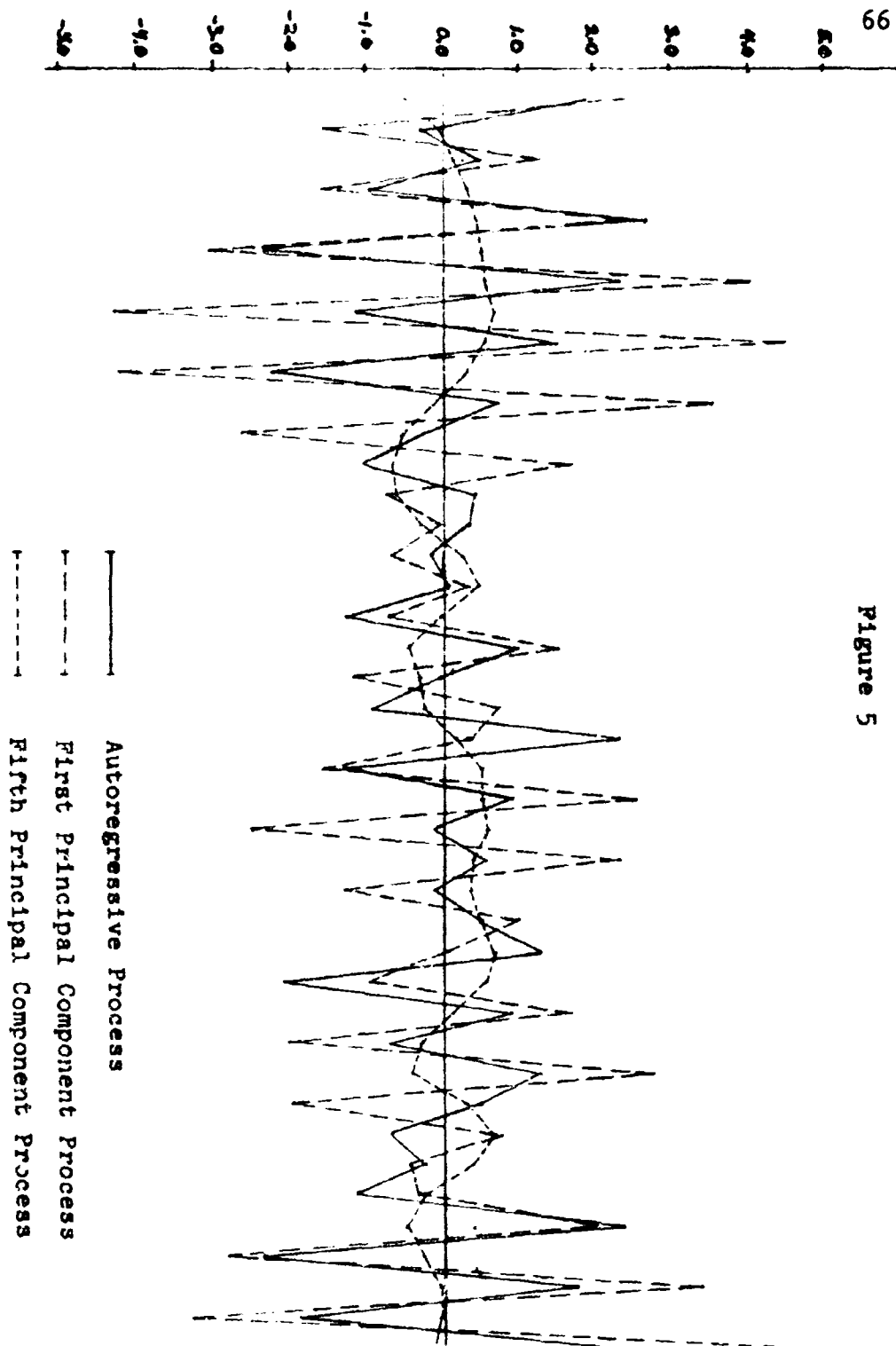principal component process contains oscillations of much
lower frequencies.

Figure 5

Autoregressive Process

First Principal Component Process

Fifth Principal Component Process

Figure 6 shows a sample of the process $x(t)$ $= \frac{1}{10} t + 5\cos\frac{\Pi}{2}t + y(t)$, where $y(t) + .9y(t - 1) = Z(t)$, $Z(t)$ a white noise process. Also shown are the first and second principal component processes for $n = 18$. The process $x(t)$ was constructed by taking as $Z(t)$, two hundred numbers from the Rand Table of Random Normal Deviates, constructing $y(t)$, and combining $y(t)$ with $\frac{1}{10}t$ and $5\cos\frac{\Pi}{2}t$. The data was keypunched and the principal components determined (from the sample covariance matrix) using the BMDOIM principal component program. This process was selected as an example since it might resemble an unseasonally adjusted economic series.

The first principal component contains the linear trend plus some low frequency oscillations, while the second principal component seems to consist almost entirely of oscillations with period four corresponding to the cosine term. The third principal component (not shown) also consists primarily of oscillations with period four as would be expected from the results of chapter III, while the fourth (also not shown) component contains oscillations of period two from the autoregressive process.

Figure 6

x(t) Process

First Principal Component Process
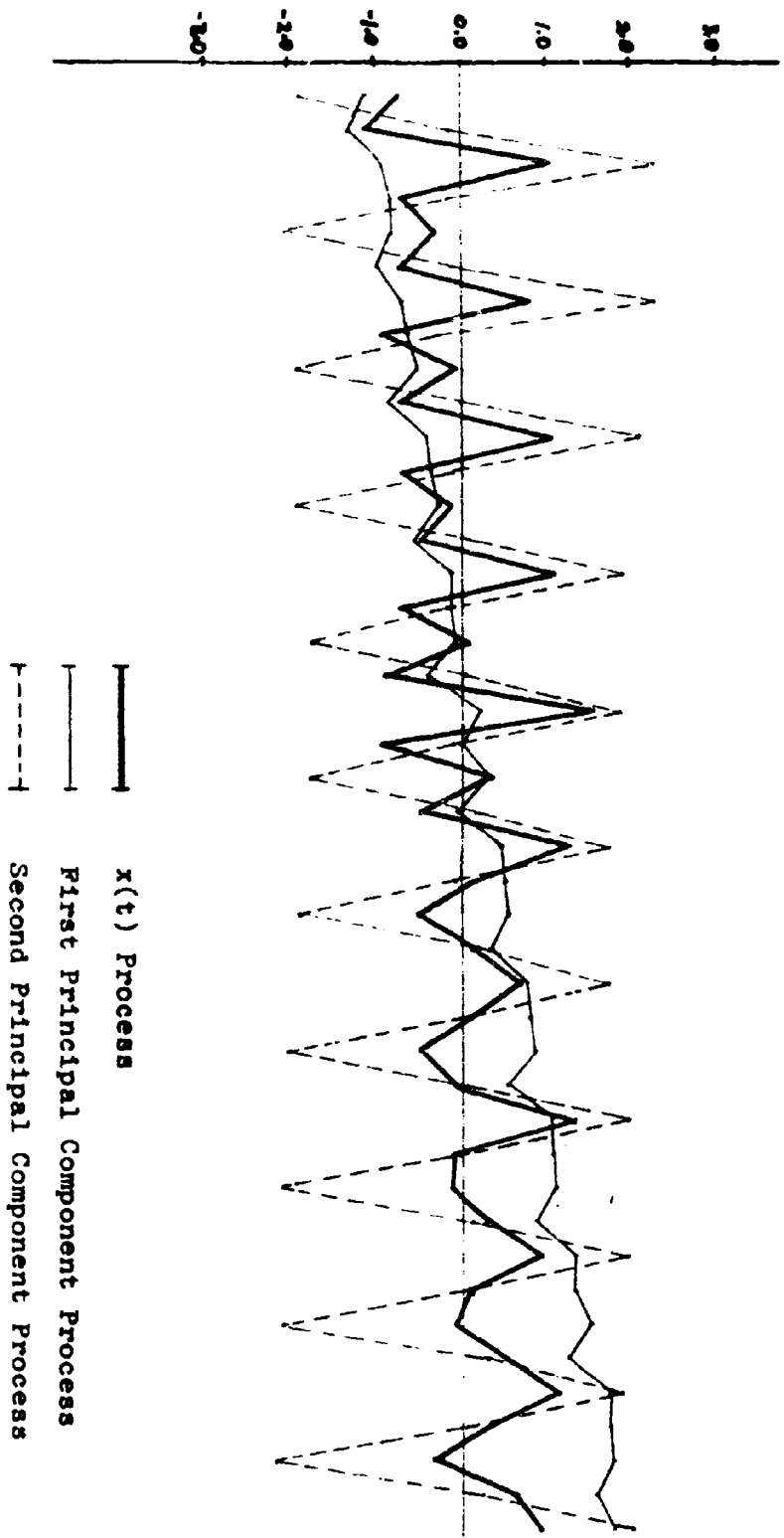
Second Principal Component Process

Figure 7 shows a series $x(t) = \log_e$ (the total U.S. Personal Income) , a quarterly, seasonally adjusted series from "Survey of Current Business", from first quarter 1958 through fourth quarter 1968. Shown with $x(t)$ is its first principal component process multiplied by a scale factor so that it fits the original series. These principal components were also computed using the BMDOIM program. In this example, n was chosen to be equal to six. Figure 8 shows the second principal component on a greatly magnified scale.

x(t) Process

First Principal Component Process
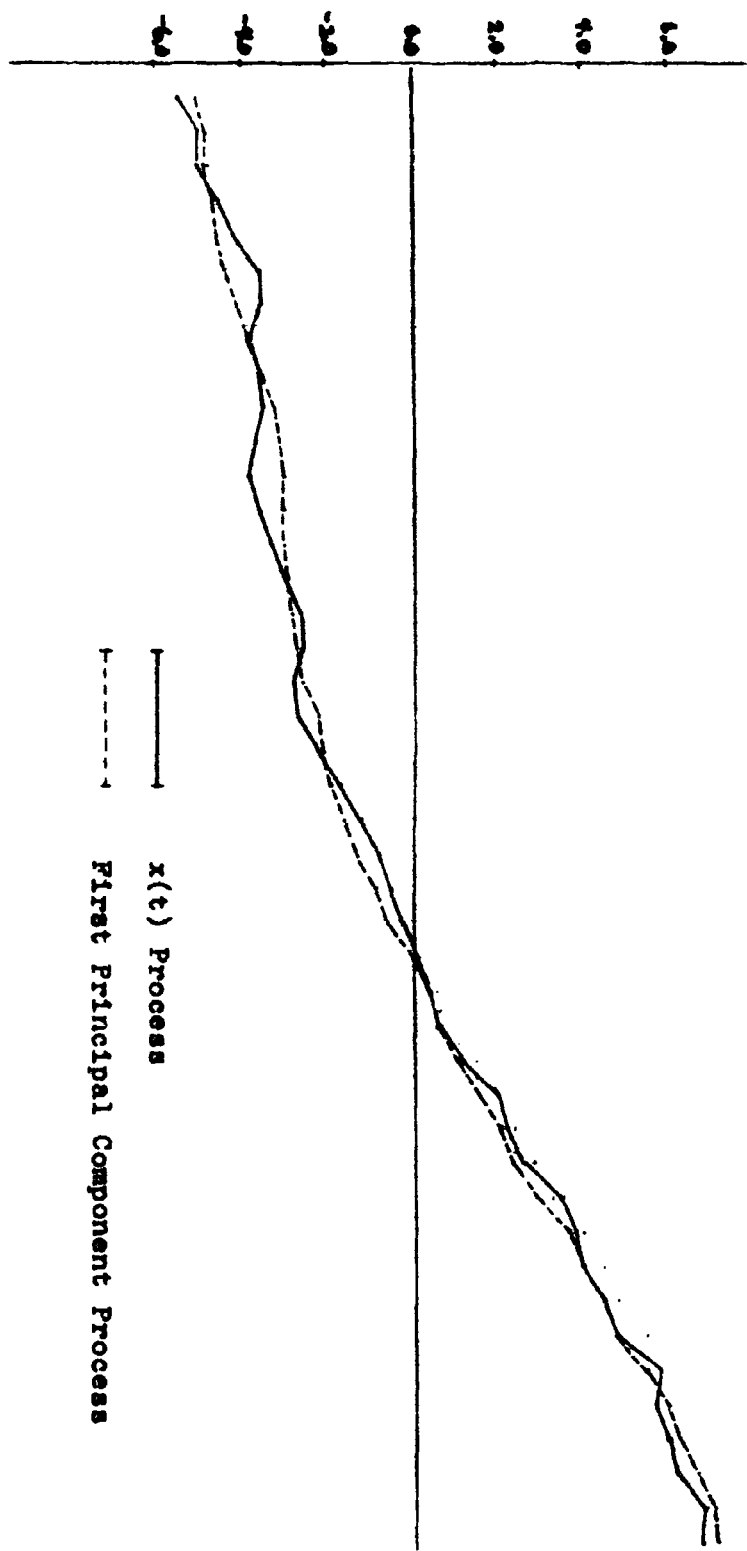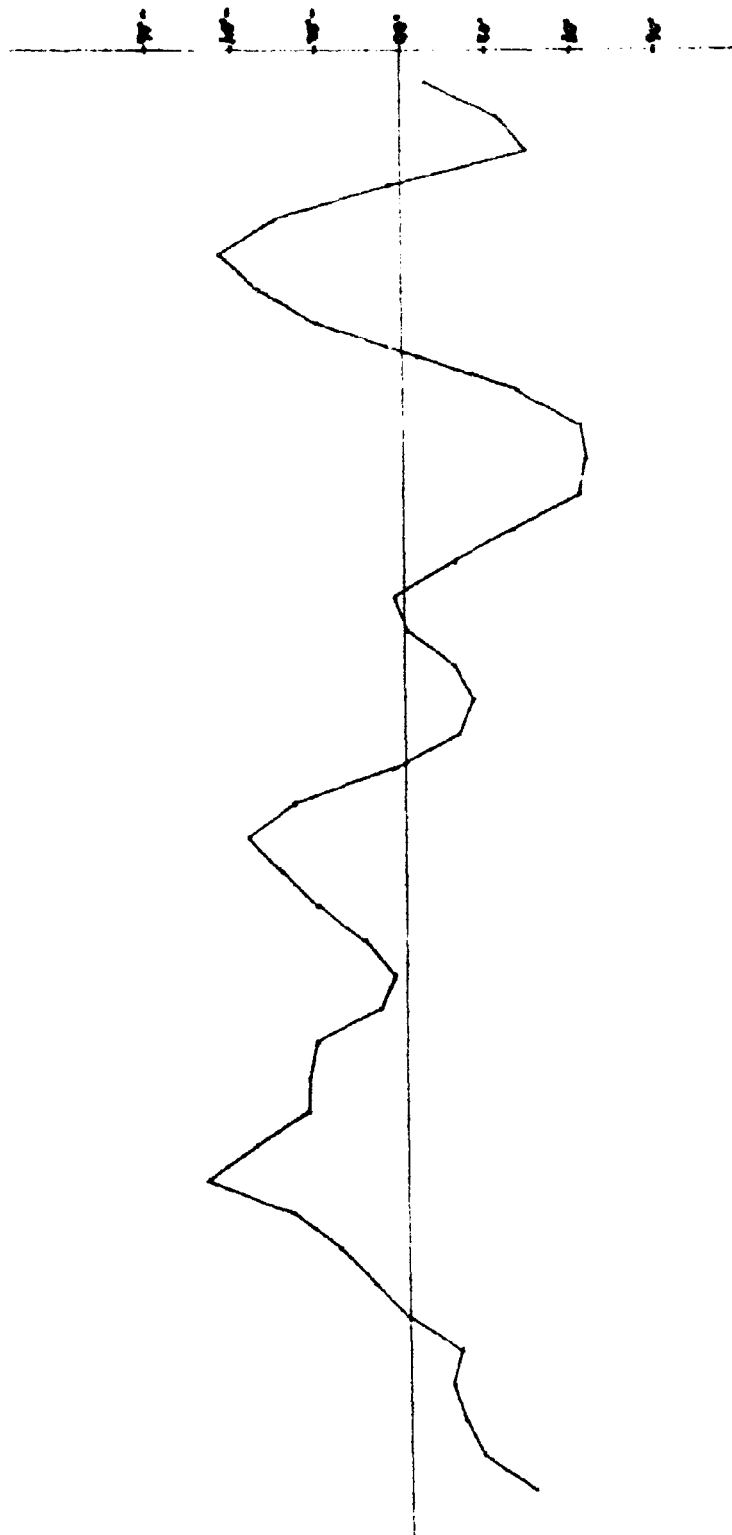
Figure 7

Figure 8

BIBLIOGRAPHY

[1]   Anderson, T.W.. An Introduction to Multivariate
       Statistical Analysis. New York and London.
       John Wiley & Sons, Inc., 1958.

[2]   Bellman, Richard.. Introduction to Matrix Analysis.
       New York, Toronoto, and London: McGraw-Hill
       Book Company, Inc., 1960.

[3]   Cramér, Harold, and Leadbetter, M.R. Stationary
       and Related Stochastic Processes. New York,
       London, and Sidney: John Wiley & Sons,
       Inc., 1967.

[4]   Franklin, Joel N. Matrix Theory. Englewood
       Cliffs, New York: Prentice Hall, Inc.,1968.

[5]   Grenander, Ulf and Szegö, Gabor. Toeplitz Forms
       and their Applications. Berkeley and Los
       Angeles: University of California Press,
       1958.

[6]   Lighthill, M.J. Introduction to Fourier Analysis
       and Generalized Functions. Cambridge:
       Cambridge University Press, 1964.

[7]   Loève, Michel. Probability Theory. Princeton,
       New Jersey, Toronto, New York, and London:
       D. Van Nostrand Company, Inc., 1963.

[8]   Massey, William F. "Principal Components Regress-
       ion in Exploratory Research," Journal of the
       American Statistical Association, March 1965.
       234-256.

[9]   Rao, C. Radhakrishna. "The Use and Interpretation
       of Principal Component Analysis in Applied
       Research," Sankhya, ser. A, vol. 26 (1964)
       329-358.

[10]   Watanabe, Satosi. <u>Knowing and Guessing. A</u>
        <u>Quantitative Study of Inference and Inform-</u>
        <u>ation.</u> New York, London, Sidney, and
        Toronto: John Wiley & Sons, Inc., 1969.

DOCUMENT CONTROL DATA - R & D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY (Corporate author)	2a. REPORT SECURITY CLASSIFICATION
The Ohio State University, Division of Statistics Columbus, Ohio 43210	UNCLASSIFIED
	2b. GROUP

3. REPORT TITLE

PRINCIPAL COMPONENT ANALYSIS OF TIME SERIES

4. DESCRIPTIVE NOTES (Type of report and inclusive dates)

Scientific   Interim

5. AUTHOR(S) (First name, middle initial, last name)

J. Richard Stewart

6. REPORT DATE	7a. TOTAL NO. OF PAGES	7b. NO. OF REFS
September 1970	73	10

8a. CONTRACT OR GRANT NO.	9a. ORIGINATOR'S REPORT NUMBER(S)
AF-AFOSR 1305-67	
b. PROJECT NO. 9749	
c. 61102F	9b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report)
d. 681304	AFOSR 70-2809TR

10. DISTRIBUTION STATEMENT

1. This document has been approved for public release and sale; its distribution is unlimited.

11. SUPPLEMENTARY NOTES	12. SPONSORING MILITARY ACTIVITY
TECH, OTHER	Air Force Office of Scientific Research (NM) j 1400 Wilson Boulevard Arlington, Virginia 22209

13. ABSTRACT

The primary purpose of this dissertation is to investigate the properties of the principal components of a finite set of random variables comprising a part of a discrete time series. In the first chapter, the covariance structure between a set of random variables $y, x_1, \ldots, x_p$, which yields the result that the first $k (< p)$ principal components of $x_1, \ldots, x_p$ provide a better predictor of y (in the sense of expected squared error) than do any k of the variables $x_1, \ldots, x_p$ themselves, is examined.

In the remaining chapters, principal component processes which are linear combinations of $x(t), x(t-1), \ldots, x(t-n)$ where $x(t)$ is a random process and n is an arbitrary positive integer, are defined and their properties investigated in terms of their frequency content. It is shown that when $x(t)$ is a stationary moving average process, an autoregressive process, or a mixed moving average autoregressive process, the first principal component process tends (as $n \to \infty$) to contain only the frequency at which the spectral density of $x(t)$ obtains its maximum value. It is shown, moreover, that when the process $x(t)$ contains deterministic components such as a trend or a periodic component, certain of the principal component processes tend to model those deterministic components.